UNCLASSIFIED

AD NUMBER
AD449586
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; JUN 1964. Other requests shall be referred to Office of Naval Research, Washington, DC.**
AUTHORITY
onr memo 6 jul 1965

THIS PAGE IS UNCLASSIFIED

MANUAL CONTROL

CHARLES R. KELLEY

PREPARED FOR:

ENGINEERING PSYCHOLOGY BRANCH
OFFICE OF NAVAL RESEARCH
WASHINGTON, D. C.

MANUAL CONTROL

Theory and Applications

Charles R. Kelley, Ph.D.

with an appendix by Meredith B. Mitchell

Prepared under Contract Nonr-4109(00)

for

Engineering Psychology Branch
Code 455
Office of Naval Research
Washington, D. C.

by

Dunlap and Associates, Inc.
Western Division
1454 Cloverfield Boulevard
Santa Monica, California

24 June 1964

PREFACE

The consultant spends most of his time on applied problems, and is likely to complain that he never gets a chance to back off and take a broader and deeper look at his area of specialization, even though he may feel his background and experience suit him well to do so. As a consultant I felt this way, and said so. The Office of Naval Research called my bluff by funding this study, which provided for a broader and deeper look at the subject of manual control than I had ever expected to be able to undertake while working in the consulting field. I'm afraid the process has been much harder than I had anticipated. I hope the end product justifies the investment, and makes it easier rather than harder for the next consultant with a similar complaint to obtain backing for a project dear to his heart.

It appeared to me that we did not really have a theory of manual control, although we did have a number of engineering models of the human operator in a tracking task. The applied problems I had encountered in the manual controlof, e.g., submarines, spacecraft and pipelines, could only rarely be described via these models, however. Usually the operator was doing something quite different from tracking. This "something different" seemed critical to the tasks performed, and yet it was just this difference that we had no adequate theory to cover.

I have endeavored here to develop and to apply a theory of manual control. It holds that the process of control begins with man and is only partially extended by control mechanisms. I have therefore started with man and the way he exercises control when there is no mechanism to aid him. It briefly discusses tools and control devices, and goes into detail in dealing with control systems and with the manual control process. The final chapters apply the theory specifically to displays and controls for the human operator. No attempt was made to cover material already adequately presented in texts and handbooks. Throughout, the focus has been the nature and importance of the role played by man in the control process.

It might be expected that a psychologist would put man in the center of a theory of control, and consider control mechanisms as secondary. Psychologists among my readers will note, however, that the roots of my theory lie, not in the behavioristic psychology of today, but in the "hormic" psychology of William McDougall, the British-American psychologist active during the early decades of the century. McDougall proposed a theory of human action based on the concept that living things organize their behavior around goals. The movements and actions of living organisms cannot be understood, he felt, except in terms of purpose, of goals, and the striving toward them.

I believe that McDougall was right. Though the words used differ from those of McDougall, the theory of control presented here is similar in essentials to McDougall's theory of action. It has the virtue -- and the weakness -- of being much more explicit than was McDougall. If there is one reservation I have about my own theory it is that it seems to imply more structure and specificity in the internal processes leading to control than is usually there. I speak of the operator's "plan", when what actually exists may be a precise pre-formed program of goal directed activity, but is often no more than a vague intention. The "operator's internal model", which plays a central role in my theory, has not the rigid structure we associate with physical models; its materials, the materials of consciousness, are fluid and evanescent, its representation highly selective and partial. It is, nonetheless, a model. McDougall wrote of human action with great sensitivity to the many undercurrents and preconditions from which an act grows. My own theory goes beyond McDougall in tracing the cause and course of goal-directed activity, and whereas I have been more explicit and, within the range of activity on which I have focussed, more complete, McDougall was more subtle and, of course, ranged through a much wider territory in his several decades of productive work.

This is a first presentation of a theory and a review and application of the theory to a field of technology. The theory will need development, clarification, and doubtless, correction. I expect soon to incorporate the material of this report in a book. I ask, therefore, that my colleagues working in manual control write me about errors of commission and omission, whether in the theory, the way it is applied, or the way I have reviewed work in the field. In particular, I would like not to misinterpret or misrepresent the work of others.

As the project progressed it became clear that there was a substantial volume of material to which I could not do justice, due to the limits of my skill in advanced engineering and mathematical techniques. My associate, Mr. Mitchell, stepped in to fill this need, and thus to round out the report. His work is incorporated in an appendix, not because it is less important than my own, but because it is not integrated with my own. It was done independently by him and thus has no reference to the theory about which the remainder of the report is organized.

My obligations in this study are many, and I can credit but a few, The late Dr. Jerome H. Ely, Vice President of Dunlap and Associates, Inc., and a close friend during all my years in the manual control field, was instrumental in my getting the opportunity to carry out this work, as also was Dr. Jack W. Dunlap. The Engineering Psychology Branch

of the Office of Naval Research provided the necessary funds and, since they have supported me also in much of my work in the past, I am once again grateful. I hope that they find this rather different sort of project one that proves its worth.

Santa Monica, California C.R.K.

June, 1964

TABLE OF CONTENTS

APPENDIX

LIST OF ILLUSTRATIONS

APPENDIX

LIST OF TABLES

APPENDIX

MANUAL CONTROL

I. THE CONTROL PROCESS

Control of the environment is perhaps the most significant of the activities of living things. It has as its necessary condition knowledge of the environment. Its motive is the desire to change the course of events in that environment. Its physical origin is movements within the body which enable the organism to locomote, to manipulate, to build or destroy, to make something happen or keep it from happening. Man's ability to control his environment is unique. It is not a product of, but is rather synonymous with, his technology. The most significant points in its history were (1) the first time tools were deliberately employed; (2) the first time external sources of energy were utilized consciously for human ends; and (3) the first time these were combined, i.e., a tool which employed an external energy source was used to control the environment. The latter is the human accomplishment from which the technology of control has grown.

The terms "control devices" and "control systems" are generally applied to only a small proportion of the apparatus and organizations by means of which man exerts control over his environment. They are a particularly interesting part, because with them the control process itself becomes the center of focus rather than the change the process is designed to bring about. The technology of control devices and control systems has this self-reflexive character; it is the technology of the control of the control process.

Manual control is a part of the technology of control devices and control systems. It can be argued that it is the most important part; consider only the role that manually operated vehicles and power tools play in our lives. The position of manual control in this technology is anomalous, however. The theory and the techniques employed by the control engineer in developing automatic control systems are not adequate for manual control. Nor is there any generally accepted theory of manual control to serve as a guide, although there is a host of useful special techniques. Since it is the aim of this study to develop a theory as well as to review existing techniques, we will begin with considerations fundamental to the control process itself. These will be formulated differently than is customary in treatments dealing with automatic control mechanisms and control devices. The aim here is improved understanding of those control systems in which man plays an integral part, and of the role that man plays in such systems.

A. Human Functions Extended by Technology

Man builds physical products which extend already present aspects of himself. Technology is based on at least four fundamental categories of such products, which can be distinguished on the basis of the human functions they extend. They are:

1. Sensing
2. Information transformation
3. Information storage
4. Control

Sensing devices or systems are concerned with the relation

$$\text{environment} \longrightarrow \text{man}$$

This unidirectional or transitive relation deals with the gathering of information from the environment. Such devices as the thermometer, telescope, or the gasoline gage in an automobile serve as extensions to the eyes, ears, and other senses. A military intelligence network or radar missile warning system serve the same function. These are classed as sensing systems rather than sensing devices only because they consist of a multiplicity of elements.

The information gathered through the senses (with their mechanical extensions) may be changed in many ways by transformational processes. It may be transduced, so that the information is carried by a different kind of energy. It may be filtered or "reduced" to eliminate unwanted information. And it may be subject to a variety of changes in form to bring out aspects of the information gathered that are not otherwise evident. Logical and mathematical transformations are one class of such changes. Computing devices, from abacus and slide rule to digital and analog computer are examples of devices which extend the human capacity for the latter class of information transformation, a major function of the brain.

Information storage, the third aspect of man that is extended by physical devices, is accomplished by imposing a special structure on some physical object, as marks on paper or magnetic patterns on tape, from which a pattern of information can be recreated for later use. Information storage extends the human capacity to remember, a second major function of the brain.

The fourth category, control, is the subject of this study. It concerns the relation

$$\text{man} \longrightarrow \text{environment}$$

This also is a unidirectional or transitive relation, depicting a flow of energy which results in a purposive modification of the course of events in the environment. Tools, control devices and control systems extend man's ability to move around in, manipulate, or otherwise modify his environment; thus they serve as extensions of the limbs and muscles.

B. How Control is Exercised by Man

1. Control and the Future

The control process is rooted in the human desire to change the future course of events. Control is of necessity directed toward the future. Past and present are immutable, beyond all possibility of modification; it is only events which have not yet occurred over which there can be control. Our knowledge of the course of events, however, is based on information from the past. We do not "have" the future, and can only infer from sensory data and past experience what the future might be and what might be done to change it. In the simplest case, human control reduces to a man knowing that if he does nothing X will happen, but if he does A, Y instead of X will result. As long as he has the choice of doing or not doing A, and can foresee Y and X as possible outcomes, he has some control over the future; he can bring about either X or Y.

Control, then, involves a choice or selection among possible future states, the chosen state comprising the chooser's goal. This choice is implicit in every control activity, whether action to achieve it is carried out by living individuals, by an automatic device or control system, or by some complex arrangement of men and equipment -- and the choice itself is always made by man. True, it can be made from a remote point in space and transmitted, or in advance in time and stored -- or a contingent choice may be made which depends on events to be detected by a sensing mechanism. Nonetheless, the decision, the choice of a goal, always originates with man. Only the conscious individual is able to conceive of different possible future states and to select from them that which he wishes to bring about. This ability alone makes it possible for man to control the course of physical events.

2. The Nature of Goals in the Control Process

The term "goal" as employed here refers to any possible future state that is selected from two or more alternatives. The goal may be ahead any length of time in the future. The president of a corporation must plan ahead for years; the driver of an automobile must plan ahead for seconds. Both must conceive of and select from possible future states as the initial stage in their respective processes of control. "Goal" is thus defined here in the most general sense. The technology of control has developed principally around more immediate goals, but the control process is the same in principle for remote goals as well.

The daily activities of men are filled with control activities. Most human activity, in fact, involves changing the environment in some way, and may be subsumed under the category "control". And human control activity is organized around the selection and pursuit of goals, short range and long.

Analysis of the goals around which particular human activities are organized shows that goals tend to be organized into hierarchical structures, with those nearer in time leading toward the more remote. The close-at-hand goals may be thought of as subgoals or routes to more distant goals. However, the conception of and choice among possible routes or subgoals is the same kind of activity as that involved in choice of the original goal. In point of fact, the remote goal may be fixed, so that the only freedom an individual has is in choice of routes or subgoals.

To illustrate: In driving to work a man's goal may be to reach the office by a particular time. This is the most remote goal around which his driving activity is organized. This goal may be determined by the circumstances of his life so that in effect he has little choice about it. He may have a choice of routes, however, particularly if the time constraints for the trip are not severe. The choice of route then involves the process of conceiving of and choosing among the "alternative future states" connected with the different possible routes. Having selected a particular route, the driver is continually selecting and pursuing still shorter range goals. Shall he drive in this lane or that? Shall he pass the car ahead or be content to follow it? How fast should he drive this stretch of road? For each such short range goal, as for each long, he must conceive of and select from alternative future states.

This simple example, which is representative of the kind of control

activities this volume is concerned with, illustrates several important points about the nature of goals:

- There may be many possible routes to the same goal.
- The choice of route is a goal selection process in itself, and the route can be considered as a subsidiary goal or subgoal.
- Even though the remote goal is fixed, there may be a number of alternative routes or subgoals to be selected from.
- Each possible route or subgoal incorporates other still shorter range goals to be conceived of and selected from, so that goals are hierarchically structured.

The hierarchical concept of control will be developed more fully in a later section of the report.

3. Goal Conception

The process by which man conceives of and selects among possible future states is the most important and least understood part of the control process. Man receives information through his senses and applies information stored in memory to create internally, from the little understood materials of consciousness, a dynamic model of the world about him. This model not only represents the structure of the environment, but also incorporates its rules of operation, e.g., temporal order and cause and effect relations. This model represents the individual's perception and understanding of his environment.

The nature of the modelling process is such that it is not limited to past and present, but can be used to create representations of possible (and impossible) future states as well. The mental activity in which possible future states of the environment are created is the goal conception stage of the control process.

In controlling a particular environmental variable, the goal conception process must take into account information about the variable to be controlled, the environmental factors affecting it, and the potentialities and limitations of whatever techniques he has available to affect this variable. When the control process is extended by a mechanical device, the "mental modelling" process must incorporate the capabilities and limitations of the device, or the device cannot be

used effectively. The ability of an individual to exercise control is limited not only by the external constraints of circumstances, but by more severe internal constraints of his knowledge.

It has been stated that control begins with the conception of and choice among possible future states or goals. This requires the individual exercising control not only to be able to conceive of different future states but also to be able to differentiate those which are possible from those which are not. Since the individual must be able to initiate a chain of events which will bring the goal about, he must therefore be aware, not only of possible future states, but also of how these states can be realized. The conception of goals in the process of control includes a knowledge of the trains of events which will lead to the goals, for it is these trains of events which make the goals "possible". The technology of control has developed around such trains of events.

4. <u>Goal Selection (Planning)</u>

The individual exercising control must choose among alternative future states, which may be discrete possibilities, a continuous range, or both. If there were no alternatives there could be no control. The choice he makes is based on the alternatives conceived of, and on their expected consequences. Increased knowledge of the alternatives available provides the individual with more possibilities to select from, while increased knowledge of the potential consequences of the alternatives permits him to make a better choice. Paradoxically, knowledge of the alternatives available widens the range of choice of the individual exercising control, while knowledge of the potential consequences has the effect of narrowing it. However, both kinds of knowledge improve the effectiveness of control by whatever criterion or criteria govern the choice of goals.

Why an individual chooses a particular alternative can be a difficult question to answer. It may appear easier, safer, cheaper, or more enjoyable than other possibilities. When a criterion is chosen, it imposes on the goal selection process a more general goal, which is at a higher level in the hierarchical structure of goals. A criterion for choosing a goal is, after all, a goal in itself that has been chosen among alternatives in a prior process of control. By imposing a fixed criterion to choose by, the range of possible choices in a control process is narrowed and may, in fact, be narrowed to one. Choice may in this way be much reduced or even eliminated at the lower level in the hierarchy of control. This does not eliminate the role in the control

process that is played by human conception of and selection among possible alternatives, but rather raises this activity to a higher level in the hierarchy of control processes, i.e., to the level of the choice of the criterion.

5. Initiating Control

The conception of possible goals in the control process requires, as was said, knowledge of the trains of events which can be initiated to bring each possible goal about; it is this knowledge that makes the goal possible. Having selected the goal, then, the train of events required to bring it about must be initiated. It is always initiated by some bodily activity on the part of the individual who conceived of and selected the goal, be it the purposive movements of his hands, arms, or larynx, or in other muscular activity. We have elected to call an individual the director of any control process which leads to a goal he conceived of and chose, and which is achieved by events he initiated. It is the unusual manual control process that does not involve some degree of goal selection, and hence direction by a human operator.

Control requires that the course of physical events be changed from what they would have been had not control been exercised. The change in the course of events requires the intervention of energy. The energetic process which brings about the change originates with the director of the control process. The director may intervene to change the course of events by means of his bodily activity alone. On the other hand, mediation of the change initiated by the director may involve tools and/or external sources of energy, or a complex and varied train of events, employing other individuals and mechanisms and spanning large distances or long periods of time. Such is the nature of the control process. In every case it has this feature, however: an energetic process triggered originally by the director's bodily activity leads to a change in the course of events that brings about a goal he selected.

The initial process by means of which the director of a control process intervenes to change the course of physical events is a mystery to the scientist, involving, as it does, the problems of "freedom of the will" and of the relation of mind and body. No attempt will be made here to deal with these problems. The common sense observation that man does change the course of events in the pursuit of goals is, however, our basic assumption.

6. Achieving Control

Control is achieved when the future state or goal that was conceived of and selected by the director of the control process has been realized. Man always initiates and sometimes completes the physical train of events leading to the goal by means of his bodily activities. The fundamental means of control of the environment is through purposive movements, i.e., movements designed to bring about the preconceived future state about which control is organized.

In the most direct form of control, the "control director" carries out the desired change in the environment himself using unaided musclepower. He pushes, turns, lifts, builds, locomotes, etc., by means of movements of limbs and digits. These movements typically consist of highly organized carefully timed precise chains of physical events. They are organized about the perception of the existing course of events and the conception of the change being introduced in striving toward the goal.

The movements by means of which control is exercised reflect the hierarchical structure of the goals of human control activities. Familiar highly practised brief patterns of movements serve the more immediate short range goals, such as standing, reaching, turning, striking, etc.; these "simple" activities are chained together to form more complex patterns of movements which are organized functionally about more distant goals. Dialing a number on the telephone, getting out a paper clip, starting an automobile, etc., involve sequences of many different "simple" movements, organized to bring about goals which may require an exact pattern and sequence of bodily movements. These more complex activities are themselves chained together in still larger patterns of activity which are organized about still more inclusive, longer range goals, e.g., telephone the druggist to refill a prescription; send a brochure and a letter to a potential client; drive home via the drug store; taxi the aircraft out for takeoff on runway 36. These in turn may be incorporated in still larger sequences serving still longer range goals, and so on, as the hierarchy of control is ascended.

When man reaches a goal, whatever level that goal may be, he has achieved control at that level. Control is achieved when a goal is reached. To summarize the total process of control:

1. The course of events is perceived.
2. Two or more possible future states and events that will lead to them are conceived of.

3. One of these future states is chosen as a goal.
4. Bodily movements are employed to initiate a train of events leading to the goal.
5. The train of events initiated by the bodily movements (which train of events is in some cases monitored and modified while in progress) brings about the goal.

This is how control is exercised by man.

C. The Mechanical Extension of the Control Process

Control is achieved by man when he has successfully altered the course of events to bring about his goal. Frequently man cannot make the desired changes in the course of events by unaided musclepower. Man possesses an ability almost unique among the animals to make use of things external to his body to expand his control over the environment. Tools, external energy sources, and their combination are the essential non-human ingredients of the technology of control. Their use expands enormously the possible changes man can make in his environment.

1. Tools

The simplest mechanical extensions of the human control function are tools. So universally are tools used by men and so rarely are they used by other species that man has been called the "tool-using animal". A tool is an object used by man to change the environment by the direct application of muscular energy. Man supplies the force which renders the tool effective. With it he may cut, pound, grind, propel an object, etc., with an effectiveness otherwise inconceivable. The tool enables man to apply his musclepower with great effectiveness.

Certain tools serve as an almost literal extension of the hand and arm. Cutting, scraping, and pounding tools illustrate this kind of extension of functions of the hand. More sophisticated tools employ human musclepower in ingenious ways in which the extension of the limbs is not so literal. This is true, for example, of tools used to throw something, like a spear, sling or sling shot, or bow and arrow. The limbs, after all, are attached to the user.

Tools include highly developed mechanisms recent in history, and new tools are still being invented. The treadle-operated sewing machine, block and tackle, hand-operated pump, and the bicycle are as much tools as are knife and axe.

2. External Energy Sources

Tools all suffer the fundamental limitation of the amount of energy man can apply via the tool. The limitation can be overcome only by tapping an external energy source. One such source is the muscle-power of other living things, animal or human. This is, in principle, the simplest use of external energy, involving just the enlargement of the principle of the muscle-operated tool to one or more other individuals. Slave or animal driven treadmills and windlasses are highly developed devices illustrating this concept.

The use of non-living sources of energy to operate tools required a major innovation in human thought. The sail capturing the energy of the wind to propel a boat may be the earliest example. It illustrates the enormous power of this innovation, for this single application of the principle greatly affected the course of human life on the planet. Subsequent early applications of the principle also captured kinetic energy occurring in nature, as in the wind or stream operated mill.

The controlled use of fire no doubt dates back to the early tool-using days of prehistoric man. The external energy source provided by fuel was applied for millenia to provide warmth, to cook, and in time to smelt metal for tools. Getting kinetic energy from fuel required another major innovation in human thought, however. It required, in addition, a highly developed technology, so that the concept antedated its practical realization by hundreds of years. The steam engine as a practical method for converting thermal energy to usable power for pumping, milling, etc., is an 18th century device, and the internal combustion engine was developed in the 19th century. Like the sail, these innovations changed the course of history, ushering in the technological revolution. The discovery of electrical energy and its application to modify the environment extended much further man's ability to utilize external energy sources to modify his environment. One of its effects was to speed the development of devices to convert one form of energy to another.

3. Powered Control

When man modifies the environment via a tool rather than by direct use of his body, the process of control is changed. Musclepower is not applied directly to the environment, but to the tool, which in turn affects the environment. Control may be more effective, but is less direct. When external power is employed for control, the process is changed again, perhaps even more radically. Man is no longer

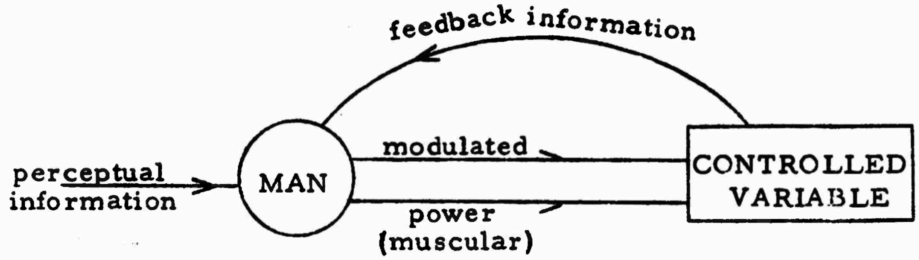
applying musclepower to the environment at all, but instead uses it to regulate the external energy which brings about the change. The relation between what the man does and what happens to the environment is rendered still more indirect. Figure I-1 illustrates these changes. The "controlled variable" is the aspect of the environment being modified in each case.

The changes are marked by the increasing indirectness of the relation of man to the variable under control, and by the increased role played by thought and the senses and the decreased role played by muscular strength. The muscles must still be relied on to produce an appropriate "control signal" (even though some external power source supplies the energy of control) and this may, in fact, be a very demanding muscular activity. Skill, however, assumes more importance in such activity than strength.

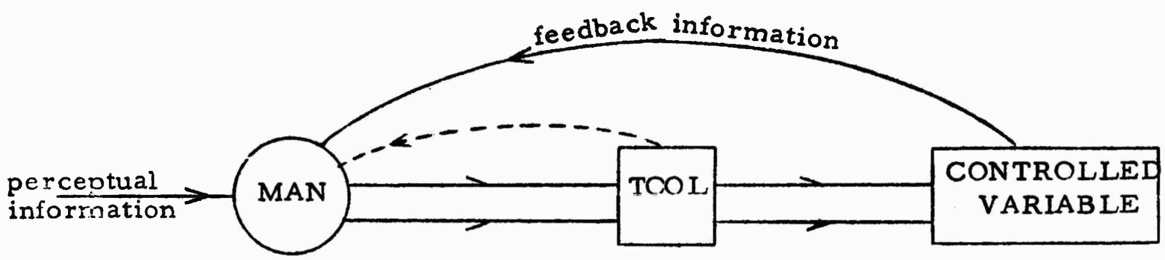
4. Automatic Control

The senses play two different roles in the control process. First, they provide perceptual information to enable the director of a control process to know about the environment and to conceive of and choose among possible future states. The senses thus aid in setting the goal of the control process. Secondly, after the goal is chosen, the senses provide feedback information to make it possible to modify and guide the activity of control. These two sensing processes are functionally discrete, even though they are going on at the same time within an individual. The sensing activity that results in goal selection leads to a man, i.e., the control system director, irrespective of how the information originates. The sensing activity providing current information about the control process, i.e., the feedback loops of Figure I-1, need not lead to a man, however. The next major development in the mechanical extension of the control process has as a characteristic feature the use of signals produced by mechanical sensors to control a source of power without human intervention.

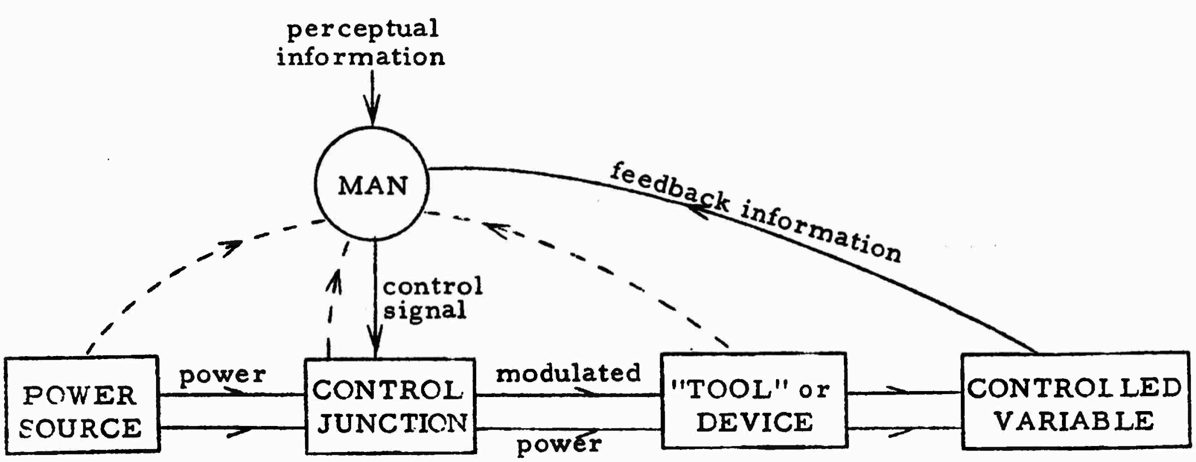
Utilization of mechanical sensors required a second development before automatic control was possible, however. The application of power to carry out control activity in pursuit of a goal depended not only on the feedback of sensory information about the course of events connected to the control process; it depended as well on information as to the goal of the control process. The controller, then, produces a signal governing the control process in accord with (1) desired, and (2) expected values of the controlled variable.



a.



b.



c.

Figure I-1. Modification of the environment via (a) direct musclepower, (b) a tool, and (c) a powered device. Dotted lines represent possible secondary feedback signals.

Historically controllers have almost always been men. Only in the last few decades has automatic control come into its own. It has been made possible by the development of both mechanized sensors and mechanized controllers, although the distinction between these was not sharply drawn in early automatic regulators. When the desired value of a controlled variable is a constant, (as is so frequently the case) its actual value seems paramount in the control loop. It is when the desired value changes with time or when two or more sensing signals are utilized in the controller that the function of the controller, as distinct from sensors, begins to become clarified It is the controller which determines how control is put into effect, the pattern of control response with time.

The development of automatic control required the isolation of functions performed by sensors and controller. Previously these functions had been confused with each other and with the process of goal selection, because they had all been performed by man. With automatic control only the choice of goal had to be carried out by man; all other roles in the control process could, at least in principle, be mechanized (see Figure I-2). At the same time, applications of the control process grew in scope and complexity, and vehicles, power tools, and other mechanisms grew in size and power. And so the control system came into its own.

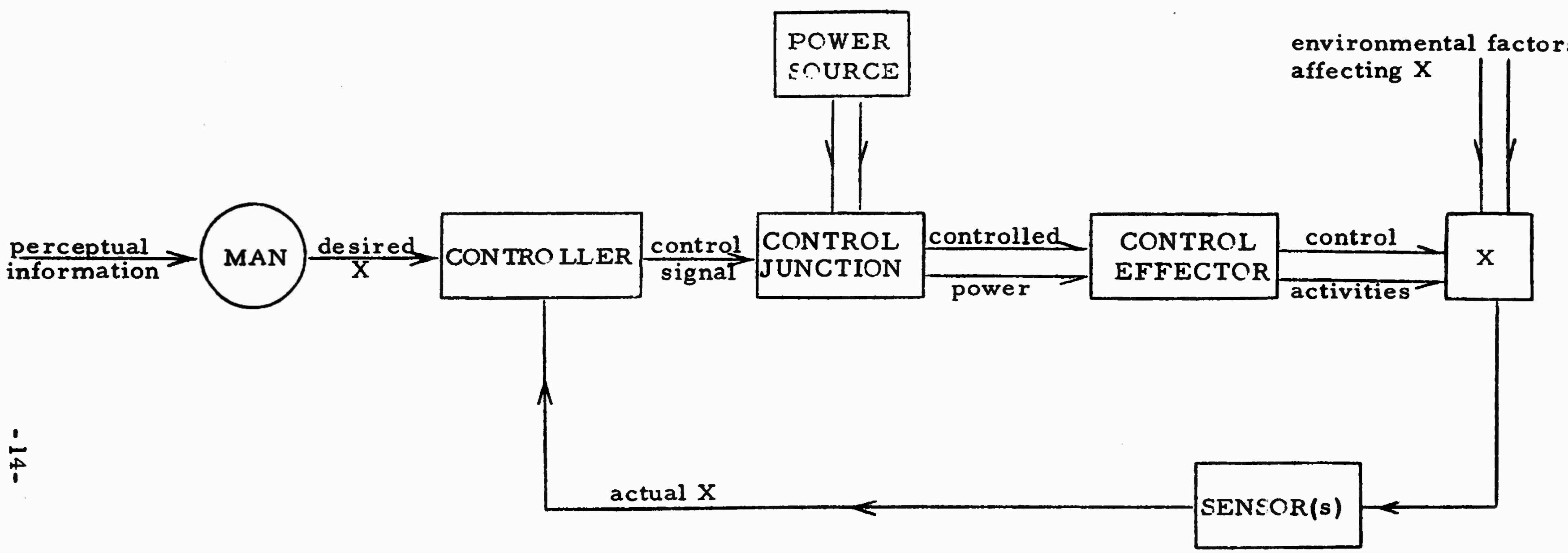


Figure I-2. Diagram of a simple automatic control system. Man supplies only the system goal or input.

II. THE CONTROL SYSTEM

The term "control system" has been used to refer to two quite different systems. Certain large military, space, and industrial organizations by their nature require a closely integrated command apparatus. Examples would be the Strategic Air Force, the North American Air Defense Command, or the ground operational support system for a manned lunar spacecraft. Systems for exercising control over such large, complex but necessarily closely unified organizations of men and equipment are, unfortunately, sometimes referred to as "control systems". The better term for these is "command and control systems". This leaves the term "control system" to designate the type of system to be described in this chapter, which consists of a single apparatus for bringing about some desired effect on the environment, albeit sometimes the single apparatus is large, with multiple human and mechanical elements. We will discuss the concept of the controlled variable, the six elements that comprise the control system, and "open loop" control.

A. The Controlled Variable

The human control process may be directed to many varied aspects of the environment; the control system, however, is by nature limited to a few at most, and usually to only one. The aspect of the environment which a particular system is designed to affect is the "controlled variable" with which that system deals. The state of the controlled variable is the control system "output". The term "output" is often used more loosely, however, as though it were synonymous with "controlled variable".

Control processes have been described as processes for changing or modifying the environment or the course of events. To avoid confusion, it should be said that one of the ways that many control systems change the course of events is to hold something constant that would otherwise fluctuate. Keeping a room at constant temperature, a vat of chemicals at constant pH or an aircraft at constant altitude are typical goals of the control process.

A controlled variable will usually be subjected to effects other than those produced by the control system. What actually happens is thus a result of a combination of effects, those stemming from the environment and those arising from the control system itself. Unless

the effects of the environment are small compared to those of the control system, changes in the controlled variable are best considered a summation or resultant of the two effects. Precise control, therefore, requires that control effects be varied to compensate for different environmental effects.

The controlled variable may be easy or difficult to define or represent quantitatively for a particular control system. Since the control process is hierarchically structured, there is a different controlled variable for each level in the hierarchy, with longer-range more inclusive "outer loop" processes incorporating the more immediate "inner loop" activities necessary to bring them about. To illustrate, consider steering an ocean liner. Depending on the hierarchical level dealt with, the controlled variable might be considered: (1) the position of the ship's rudder; (2) the direction of the ship in the ocean; or (3) the position of the ship with time in the ocean. Control systems for achieving these successively more general controlled variables might be called: (1) the rudder positioning system; (2) the ship steering system; and (3) the navigation system, respectively. Each successively more general system incorporates the preceding system(s).

When analysis deals with one of the more limited systems, the more general system is treated as part of the "environment" that is being affected by the control process. Thus, the ship's rudder is the "controlled variable in the environment" that is affected by the helm wheel and hydraulic motor of the rudder positioning system. In considering the ship steering system, however, the rudder is part of the control system, the "controlled variable in the environment" being the ship's instantaneous direction of motion. The entire ship is within the navigation system, however, the controlled variable being the location of the ship with time.

B. Description of the Control System

The control process has been described as the purposive modification of the environment. Originating as the typical activity of living things, it has been extended in man by tools, the use of external energy, control devices and control systems. A "control device" is a mechanism which utilizes an external source of power for the purposive modification of the environment. A control device is also a simple control system. The term "control system" applies as well to more complex or extensive arrangements of human and mechanical elements which have the same functional relations as do the essential parts of control devices. The

essential elements of the control system are the following:

1. Goal Selection (Planning) System: That human element or group of elements centering about a man who decides the desired modification in the environment that the system will operate to achieve.

2. Controller: The device which produces a signal to the control "junction" to release or modulate the energies of control as a function of information from both the goal selection system and the feedback sensor(s).

3. Power Source: The source of the energy of control.

4. Control Junction: The junction at which energy from the power source is released or its flow regulated in accordance with a signal from the controller.

5. Control "Effector": The element of the system applying energy from the power source to modify the environment; the major active part of the usual control system.

6. Feedback Sensor(s): An element or elements transmitting to the controller information about the aspect of the environment being modified.

Engineering descriptions of control systems omit the goal selection stage, beginning with an input signal which represents the goal. For purposes of the present study, the inclusion of this stage of the process is essential. (See Figure II-1.) The remainder of the control system is regarded as an extension of the goal selector's means for controlling the environment, while the human element within the control system typically functions by selecting and achieving subgoals at his level of the control system hierarchy.

1. Goal Selection (Planning) System

The conception of and choice among possible future states, and the initiation of a train of events to bring about the selected state or goal is the function of the "goal selection system". That it is a fundamentally human process is evident; mechanisms have no power of conception. However, the person who selects the goal of the control system may be aided in his choice by non-human extensions of the processes by which he perceives the environment and stores, makes transformations in, and uses information.

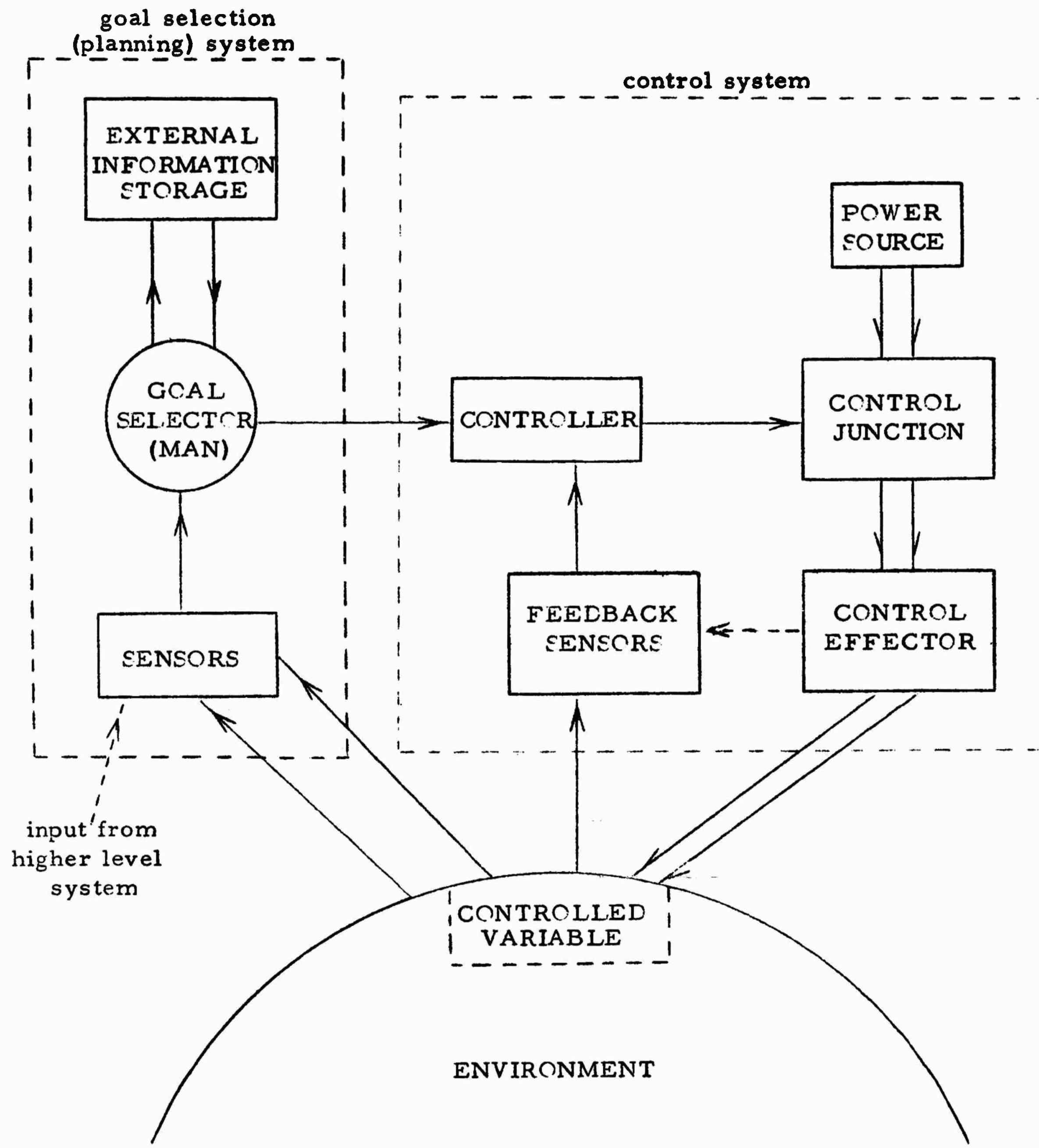


Figure II-1. Control system, including goal selection elements.

Man does more than sense the environment, he perceives and understands it. To repeat, from the poorly understood content of consciousness, he generates a dynamic model of the environment. This model is affected but by no means determined by information from past and present about the environment, not only with reference to appearance, but with respect to the environment's basic structure, including its response to the application of energies of control. By manipulating this internal model and through it foreseeing the results of various possible applications of the energies of control, man conceives of possible future states, one of which is chosen as a goal.

The input to the goal selection system is the information, present and past, that permits certain of the various possible future states to be imagined. The output of the system is a signal designed to bring about one of these states. The output of the goal selection system forms the input to the controller and thus to the remainder of the system. The remaining components of the control system can in principle all be mechanized. Treatments of automatic control customarily begin with an input signal representing the system goal as "given", and proceed from there.

2. Controller

The controller is the element which produces the "control signal" to release or modulate the energies of control. It does this in response to an input from the goal selection system which represents the desired value or state of a controlled variable, plus (ordinarily) one or more sensing signals representing the actual value or state of the controlled variable. The special case of open loop control will be discussed later.

The controller is historically the last control system element that has been mechanized. Previously, the control signal was of necessity produced by man, and its production was frequently the most difficult part of the job of control. The separation of the function of the controller from that of goal selection and sensing has made it possible to automate the controller's function entirely -- and this in turn made it possible to use man in different roles in the control system. No longer required to produce the control signal, he could be employed in various other roles in the system -- as just a sensor, for example, or to perform transformations on signals, or to serve as an "adaptive" element adjusting an automatic controller.

The controller has at least two inputs, i.e., signals representing desired and actual states of the environment or values of a controlled

variable. It has a single output, however, the control signal which operates the control junction. This signal is usually amplified in some way before it reaches the control junction.

3. Power Source

The power source forms the source of energy employed in the control system to modify the environment. Examples would be electric power or fuel capable of driving a turbine or internal combustion engine. Sometimes a power conversion is made, as when a control system is operated by hydraulic or gas pressure, and fuel or electricity is employed to operate a compressor or hydraulic motor to maintain fluid or gas pressure. Whatever conversions of this kind may be employed, the power source of the control system as defined here is that source of the energy of control that is present at the control junction and released or modulated by the control signal. The input to the power source is fuel, electricity or other source of energy, the output the power, whatever its form, transmitted through the control junction.

4. Control Junction

"Control junction" is perhaps a poor term for this element of the control system. Control "valve" better suggests the function involved, but is overly specific, implying as it does a particular kind of junction. The control junction is here defined as a device by means of which energy is released or regulated as a function of a control signal. The energy which brings about the modification of the environment is normally large compared with the energy present in the control signal. "Control junction" is thus employed here as a generic term for devices for the continuous control of large amounts of energy by small. A few examples may help clarify the function, which plays a central role in every control activity.

a. The pre-eminent example of a control junction is biological, the production of animal movement. In response to a weak neural message (control signal) muscular activity results which leads to controlled movements involving relatively high energy expenditures. The power source is the metabolic energy stored within the muscles.

b. More commonplace to the engineer is the hydraulic valve. The adjustment of a valve aperture regulates the rate of flow of fluid maintained under pressure behind the aperture which forms the power source. The fluid flowing through the valve supplies the energy to the control process, and the adjustment of the aperture via the valve is the control signal.

c. In the internal combustion engine of, for example, an automobile, it is the rate of flow of fuel into the carburetor, regulated by a control signal relayed via the accelerator from the driver's foot, that supplies controlled energy to the engine.

d. The traditional electrical example of the control junction is, of course, the vacuum tube triode, with the weak control signal on the grid regulating the much more powerful current flowing from cathode to plate.

In many control systems the control signal goes through several stages before the control junction is reached. For example, a human operator may produce the control signal by positioning a control which varies an electrical voltage via a potentiometer; this may be converted to alternating current and amplified to position a servomotor; the servomotor may then operate an hydraulic valve, which serves as the control junction proper. (This is, in fact, a fairly typical example.) Intervening stages may occur between any of the elements of the control system, of course, but are most likely between controller and control junction as part of the transition from the low power output of the controller to the high power output of the control junction.

The control junction has two inputs, an unregulated input of power or available energy that is usually large in energy magnitude compared with the second input, the control signal. The control signal is usually a relatively weak signal that regulates the flow of energy through the junction. The output of the control junction is the regulated power which operates the "control effector".

5. <u>Control Effector</u>

"Control effector" like "control junction" is a somewhat awkward expression. It is borrowed from the psychological literature, where "receptors and effectors" refer, respectively, to an organism's means for receiving information from, and for affecting changes in, the environment. The control effector is here defined as the control system element which applies the energy fed through the control junction to bring about the desired modification of the environment. The control effector is likely to involve a major part of the physical structure of the system. The effector portion of an automobile would include engine, motor drive-train, wheels and chassis, all of the parts in which the energy generated by gasoline combustion is channeled to move the vehicle. In a control system consisting of a power tool, the control effector is essentially the tool itself, (e.g., the saw blade or the drill) and the structure required to make use of the tool.

The input to the control effector is the modulated or regulated power from the control junction that operates it. Its output is the effect on the controlled variable which is the raison d'etre of the entire control system. As indicated in discussing the controlled variable, what is control effector at one hierarchical level may be considered part of the environment at a lower level in the hierarchy.

6. Feedback Sensors

Feedback sensors are distinguished from the sensors providing information to the goal selection system. Feedback sensors have as their function helping the system achieve an already selected goal by providing information about the variable under control, and (sometimes) about other information related to this variable. In open loop control, feedback sensors are absent entirely, and there is no feedback loop. This distinction is elaborated below in the discussion of open loop control.

The input to the feedback sensor(s) is information about the controlled variable. This information may be obtained through a human sense organ, through a sensing device appropriately responsive to an aspect of the environment, or through a sensing system, e.g., a radar system. The sensing system might contain both mechanical and human elements. The output from the feedback sensor(s) is a signal to the controller carrying information about the controlled variable. This may include information about factors in the environment affecting the controlled variable and/or information from the control effector itself.

C. Open Loop Control

The above six elements are found in most but not all control systems. One class of system in which they are not all present is the open loop control system. Here there are no feedback sensors, human or mechanical. In fact, the best definition of open loop control is control in which feedback information is absent.

All control is of necessity made with reference to the environment, and all control systems are of necessity governed in part by sensing information. Knowledge of the environment is a necessary condition for control. In the case of open loop control, however, the knowledge is gathered prior to the control activity by the sensing process that feeds the "goal selection system". In open loop control, there is no means for modifying control activity in progress as a function of current

information from the environment. The control activity is pre-programmed, and runs its course irrespective of what is happening.

Some discussions of open loop control appear to class systems as open loop in which feedback information is supplied by man. Such a usage is confusing. If feedback information from any source can modify control in progress, the control process should be classed as "closed loop". The confusion results from the failure to include as part of the control system functions that are performed by a human operator.

Open loop control systems are actually a rather special case, and most manual control problems involve closed loop systems. However, the human operator does behave for brief periods in a pre-programmed "open loop" fashion in certain continuous control tasks, while such common discontinuous manual control problems as firing a bullet or other ballistic projectile can be classed as open loop tasks.

III. THE CONTROL SYSTEM HIERARCHY

As the goals of a control process are organized into subgoals and sub-subgoals, so the control systems designed to achieve these goals are organized into subsystems and sub-subsystems. The very definition of the control system depends on the hierarchical level under consideration, for what is part of a control system at one level is part of the environment being controlled at a lower level in the hierarchy.

A. The General Control System Hierarchy

The control system hierarchy comprises chains of events that are utilized in the process of control, with smaller, more immediate events which employ less energy serving to bring about larger, more distant events involving more energy. The control system is built about a chain of cause and effect, and it is by utilization of this chain that small forces are able to bring about large and significant effects. Each variable in the chain of cause and effect utilized by the control system might be considered as the output of one level or loop in the hierarchy of control, the progression being from inner to outer loop as the hierarchy of control is ascended. To illustrate, consider the control of an environmental variable X, where X is brought about by Y, and Y is brought about by Z. The energy of control is applied to Z in the innermost loop. Z is varied to bring about a desired change in Y which will in turn bring about the desired change in X in the outer loop. Figure III-1 diagrams the relationship. Separate control systems might be involved to control each of the three variables, the outputs of the outer loop controllers forming the inputs of the controller for the next loop in, or the same control system might be employed to control all three. It is important to note that the desired state or value of an outer loop variable does not usually specify the inner loop variable that brings it about, but only sets limits or constraints on it. There are usually not one but many routes to a goal, not one but a range of outputs that will satisfy the requirements of control.

To illustrate the concept of control hierarchy by means of a more concrete example, consider once more the control task involved in bringing a ship from one location to another. The hierarchical structure of the task can be described in terms of goals and subgoals, the first or highest of which comes from outside the ship control system from some still higher hierarchical level:

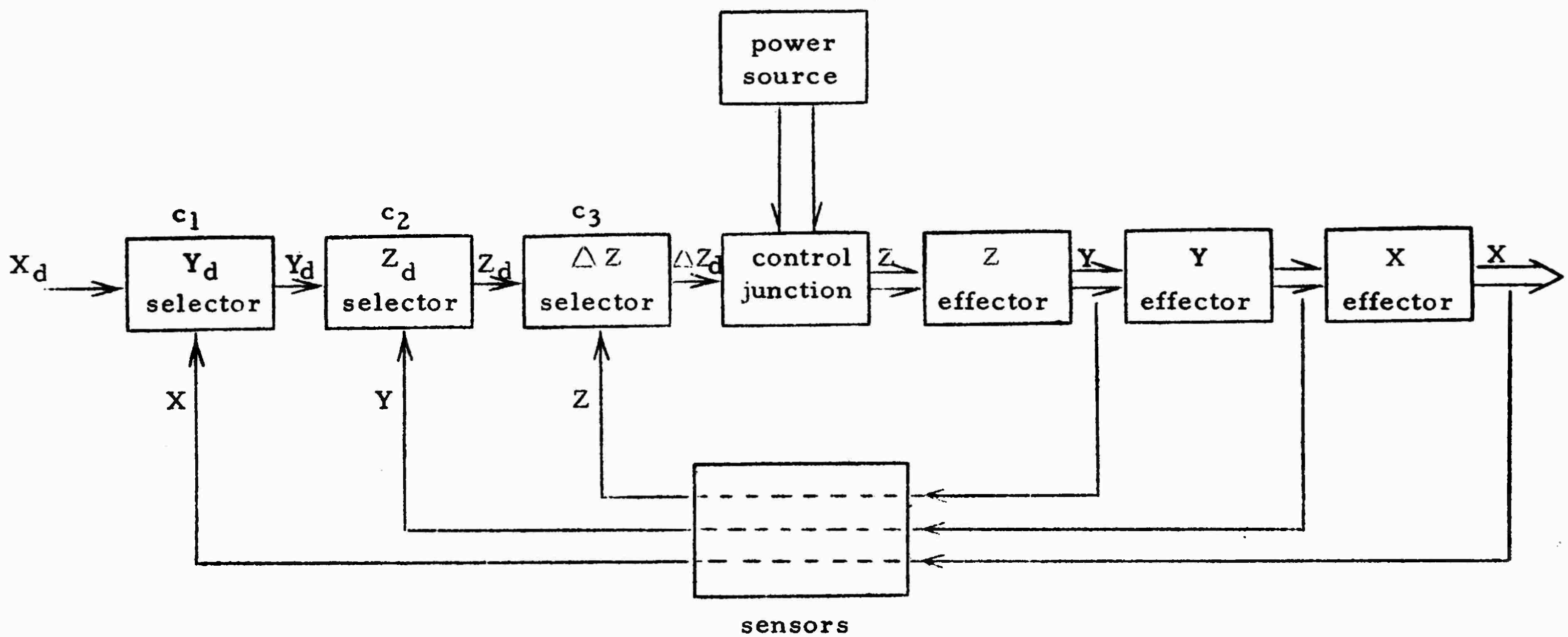


Figure III-1. General hierarchical system to control X when X is controlled by changes in Y, and Y is controlled by changes in Z.

1. Carry passengers and freight from point A to point B within constraints of schedule.

   a. Apply thrust along ship's axis via screws.

      (1) Call in desired speed from helm station to engine room via engine order telegraph.

         (a) Operate engine throttle controls.

   b. Maintain appropriate ship heading in face of disturbances as ship is underway.

      (1) Adjust rudder angle via helm wheel to apply appropriate moments to ship to maintain desired heading.

         (a) Open and close hydraulic valves to move rudder to position desired.

Each control task at a given level includes the lower level tasks necessary to carry it out. Thus it is appropriate to speak of the longer range tasks as outer loop tasks, which incorporate within them inner loop tasks necessary for their completion. The inner loop tasks themselves may contain within them inner loop tasks necessary for their completion, etc. The above hierarchy of tasks can thus be diagrammed as in Figure III-2, which illustrates the ship steering portion of the hierarchy above.

The relation of outer to inner loops defines the structure of the control task each loop represents. The outer to inner loop relation is one of goal selection or planning; the outer loop forms the means by which the operations required by the inner loop are specified. The inner loop to outer loop relation is one of cause and effect; the inner loop forms the means by which a goal established in the outer loop is reached. Inner loop tasks generally involve smaller elements of the environment that change more quickly and more frequently than the outer loop changes they bring about.

B. <u>Control Order</u>

Not only are inner loop processes generally smaller, higher frequency, etc. than outer loop; they also bring about only rates of change or accelerations in the outer loop variables they affect. Thus the <u>position</u> of a hydraulic valve in the rudder control system (inner loop of Figure III-2) results in a <u>rate</u> of movement of the rudder in the

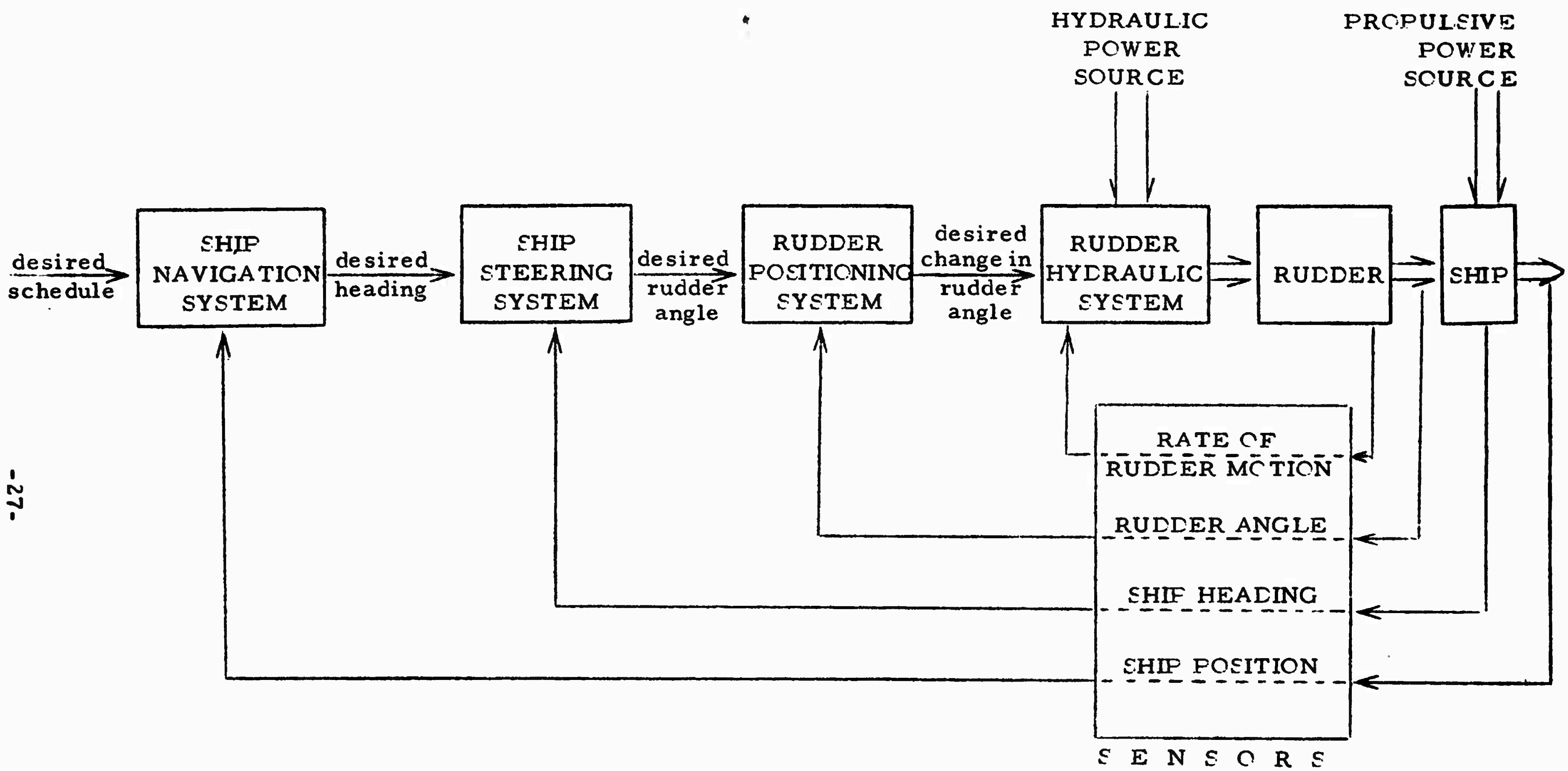


Figure III-2. Hierarchy of tasks involved in ship steering showing relation of inner to outer loops. Speed control could be represented in similar fashion.

next loop out; in this next loop, the position of the rudder results in an angular acceleration of the moving ship; the angular position (heading) of the moving ship results in a rate of change of lateral position with respect to the desired course in the outermost loop. In this typical example, progression from inner to outer loops involves not only larger more slowly changing elements, but also the progression from derivative to integral functions. The chains of cause and effect in continuous control processes are not usually of the form "Z causes Y which causes X", but rather, e.g., "Z changes the velocity of Y which changes the acceleration of X". Expressed more generally, it takes the form, "Z changes the Nth derivative of Y which changes the Mth derivative of X". To the extent that hierarchical relations between inner and outer loops in control processes correspond to the mathematical relation of derivative to integral, the hierarchy of control can be described more formally in terms of control order. Higher order terms are higher derivative terms, and represent the inner loops of the control system hierarchy.

When a concept such as that of control order is applied to real physical processes, the application frequently becomes complex, although the general principle is simple. All control systems modify certain aspects of the environment through the expenditure of energy. The modification takes place according to processes which can usually be described in terms of known physical laws. Control of the position of objects having mass can be described using Newton's laws of motion, while heat exchange, chemical processes, electrical and magnetic effects, etc., may be involved in other kinds of control processes. Physical science has provided techniques for describing changes in each of these categories that may be brought about by control systems. The most elegant and precise descriptive tool is the differential equation. The differential equation is, by its nature, structured in terms of "order".

It is instructive to examine some effects of control order in itself. Figure III-3 shows some of these graphically. The smoothest kind of change in a controlled variable would, if achieved by changing a fourth derivative function of that variable, require as an absolute minimum five changes in the direction of motion of a control. To apply this to our previous example, if the upper (zero order) curve corresponds to the lateral position of a ship with respect to a path in the ocean, the first order curve represents ship's heading, the second order curve rate of change of heading, the third order curve rudder angle, and the fourth order curve rate of change of rudder angle, which on large ships may correspond to helm wheel (and hydraulic valve)

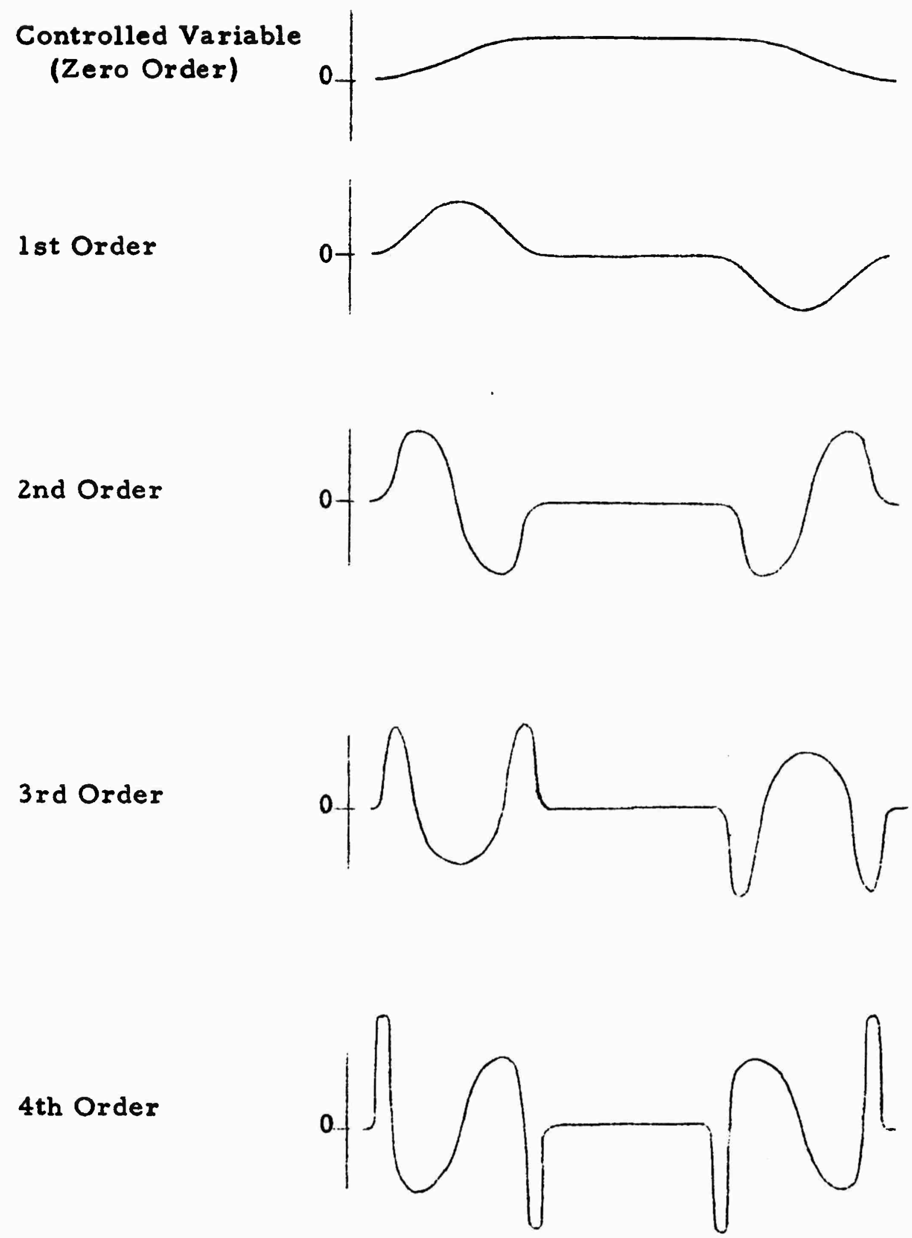


Figure III-3. Control Order illustrated by changes in a controlled variable and in its first four derivatives.

position. The higher order inner loop functions are of necessity higher frequency. (There is a mathematical exception to this rule. The derivative of a sine wave is a sine wave of the same frequency, so that if the upper curve of Figure III-3 were varying sinusoidally, the curves below would have the same frequency. Practically speaking, however, frequency increases with control order.)

Any high order control system can be structured hierarchically with successive inner loops corresponding to derivatives of the controlled variable. This control order hierarchy is a particular and particularly important form of the generalized hierarchy of control. Figure III-4 shows an idealized third order control system structured in this way. The three integrations diagrammed on the right represent the physical law relating the output of the control junction and the output of the system controlled. The computing operations performed in the controller are here broken into stages by control order, there being a zero order stage plus a stage for each integration of the control junction output, or a total of four stages for a third order system. The boxes labelled 0, 1, 2, and 3 show the signals required at each stage in the controller for a simple, smooth change in output, as previously illustrated in Figure III-3. (Figure III-4 should be compared with Figure III-2 to see how these functions compare with the actual case of ship steering.) Beginning with the output of the goal selection system, operations in Figure III-4 may be performed mechanically or by man or in combination. In particular, man may assume but a part of the operations of the controller. When this is the case, the part that he assumes is highly significant.

C. Control vs. Display Augmentation

If the human operator who is sharing the function of the controller with mechanical elements produces the controller output, i.e., if he himself operates the control junction, he is necessarily tied to the inner loop of the multiloop system as shown in Figure III-4. The mechanical elements perform one or more of the outer loop controller functions, operating on signals before they reach the human operator and displaying the computed result to him. Such systems are said to have an "aided", "augmented", or "quickened" display. When man's muscular strength is employed directly at the control junction to turn a steering wheel, position a control surface, or open and close an hydraulic valve, it is natural to turn to operations on man's input, i.e., to improve his display, in order to simplify manual control. Display augmentation is one of the several major techniques available for improving a display.

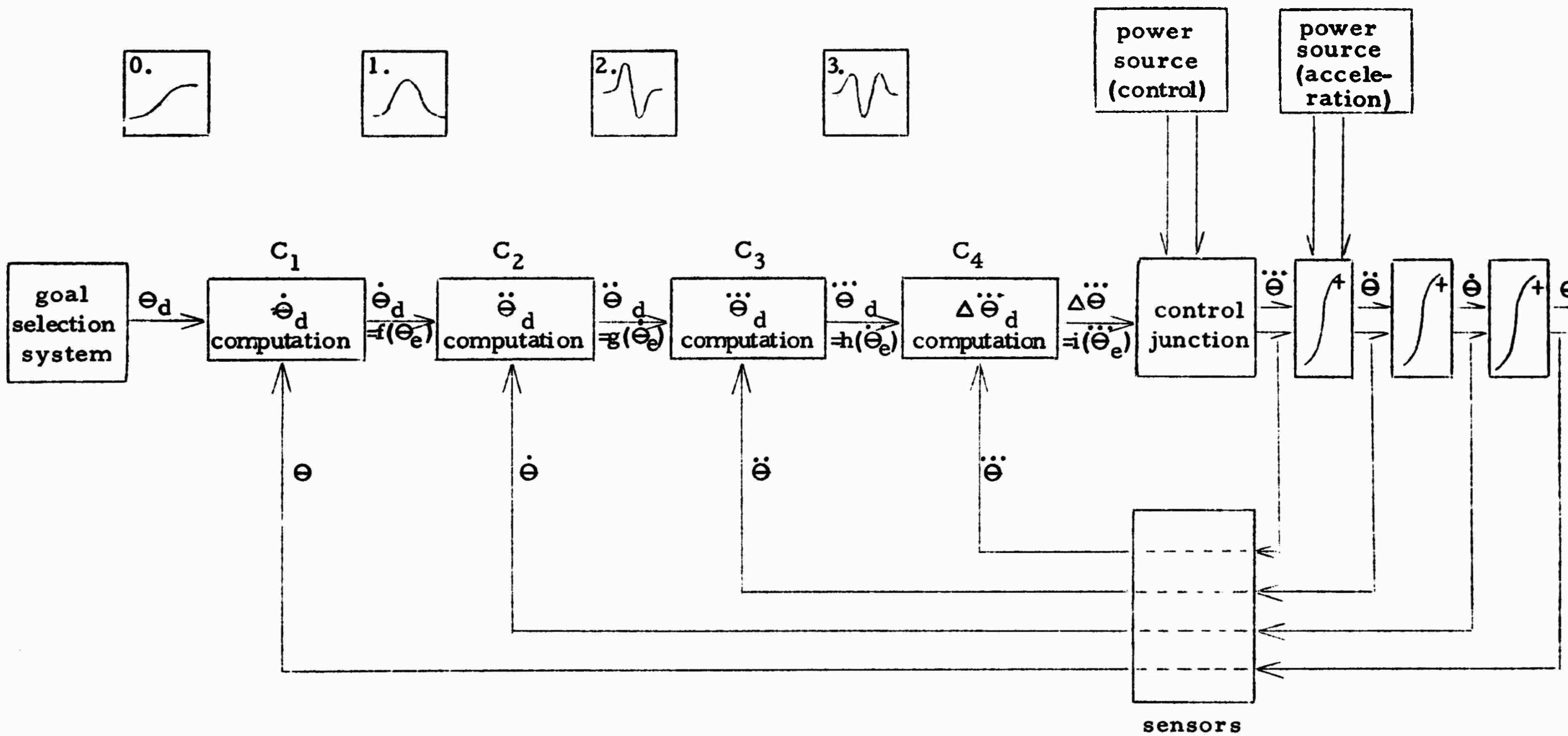


Figure III-4. Idealized third order control system, arranged to show control order hierarchy.

If man senses the output of the multiloop control system directly rather than via mechanical sensors, this too serves to fix his position in the system, but this time in the outer loop as shown in Figure III-4. Mechanical elements sharing the controller function with him may, in this case, operate on the signals produced by man and by inner loop sensors, perform the necessary computation on them, and produce the control signal mechanically at the amplification necessary to operate the control junction. Systems performing such operations on the output of a human operator are said to have "augmented control", or in vehicle applications, this type of arrangement is sometimes referred to as an "augmented vehicle". Stability augmented aircraft, in which there is an inner control loop independent of the operator, are an example.

In terms of Figure III-4, the completely manual system would have the operator perform all of the sensor and controller functions. In display augmentation, however, one or more of the outer loop sensing and controller functions would be performed mechanically, and the result displayed to the operator. The operator would perform the remaining operations, including a final amplification of the control signal at the control junction. The time structure of the operator's output for a smooth change in system output would still be that of the highest frequency and most complex signal of the series of four shown in Figure III-4.

By contrast, in control (or vehicle) augmentation, one or more of the inner loop sensing and controller functions of Figure III-4 (including any final amplification of the control signal prior to the control junction) are performed by equipment after the human operator performs his functions. The time structure of the operator's output for a smooth change in system output is, in this case, one of the lower frequency, less complex signals of the series shown. Which one it is depends upon the amount of control augmentation that is employed.

When all of the controller functions are performed automatically but the human operator is retained in the loop, it may be with a completely augmented display or completely augmented control, depending on whether he is retained in the inner or outer loop. If it is in the inner loop, he serves as a simple amplifier of the controller output. Such systems are sometimes called "semi-automatic" or "fully quickened". The time structure of the operator's response is, of course, the more complex inner loop one.

In the outer loop or augmented control case, the operator serves to sense the desired and/or actual output, either directly or via a

sensing instrument, and to enter a signal representing it (them) into the controller. In this case, the operator's response is the less complex outer loop one. Some tracking systems are of this variety.

The two roles of the human operator have these important differences:

Display Augmentation	Control Augmentation
Can employ man's muscular strength at the control junction to move a control surface, open an hydraulic valve, etc.	Can employ man's senses to observe the controlled variable and the environment.
Man is kept aware of control signal.	Man is kept aware of system output.
Requires higher frequency, more complex human output.	Requires simpler, lower frequency output.

These characteristics will often dictate which of the two methods of splitting control system functions between man and equipment are best in a given case.

It should be pointed out that control and display augmentation are by no means the only ways of simplifying the task of a control system operator. In particular, it is often possible to leave all the controller functions to be carried out by a human operator, i.e., rather than to assign either inner or outer loop functions to equipment, to simplify the manual control task by displays. Display simplification may take other forms than display augmentation as defined above, as Chapters VI and VII will indicate.

D. The Hierarchical Evolution of the Control Process

As the operator becomes skilled in control involving a number of levels or loops, he pays more and more attention to the outer loop processes, while the inner loop processes become increasingly automatic. The student driver is aware of what he is doing with his hands and feet, but as he learns, he will focus on where the car is going. Having decided where he wishes to drive, the motor processes which achieve the goal will be triggered and run their course. In the course of training, the control process evolves in the direction of

increasingly automatic performance of the inner loop activity, which makes it possible for the operator to attend more fully to the goal selection process further out or up in the hierarchy of control.

As control systems involving men and equipment evolve, it is the contention of this author that they, too, should naturally tend to develop in the direction of automation from the inner loops out. If any controller functions are to be automated, it should usually be the inner ones. Control augmentation is normally to be preferred over display augmentation. The planning capabilities of a human operator should usually receive precedence over considerations of his muscular strength and motor skill, even though the latter must continue to play a key role in many manual control systems.

IV. THE HUMAN OPERATOR: 1. CONSCIOUSNESS AND PLANNING IN MANUAL CONTROL

Human control activities have two aspects, planning (as we shall now call goal selection) and motor performance. Planning activities are unique to man, stemming from the fact that man, unlike any mechanism, is able to perceive and understand the environment, and to choose how he wishes to modify it through the control process. The motor performance aspect begins where the planning aspect leaves off. When the operator has planned what to do, he must put it into effect by means of bodily movements. In terms of the hierarchy of control just discussed, planning activities are "outer loop" activities, while motor performance deals with the "inner loop" of the human control process. The present chapter discusses the planning stage of manual control, while that which follows deals with motor performance.

Man's ability to control the environment stems from his ability to perceive and to understand it, to observe the operation of cause and effect, to predict what might be the consequence of this or that activity. All control processes, manual or automatic, derive ultimately from such conscious processes. There has always been controversy about conscious processes, for they are poorly understood. In particular, the way in which conscious processes relate to the rest of nature is a subject that occasions dispute. The basic assumption followed here is that an individual's conscious processes (such as perceiving and understanding the environment and planning to change it) do bring about the physical behavior that forms or initiates the control process. The control process originates, not with the observed physical behavior, but with the conscious processes which give rise to it. The nature of such processes is therefore of concern.

A. The Nature of Conscious Processes

There is an immense difference between physical nature and its representation in consciousness. Physical nature is an organization or structure of energy and nothing more; it is a world of geometry. The familiar world of objects, people and events is a mental creation, an internal model we build that represents a selected part of the physical world outside. It is related to the physical world only as a model is related to the object modelled. To understand it we must understand something about how the modelling process works.

1. Consciousness as a Natural Unit

Perhaps a most notable first fact about the world of consciousness is that each individual has his own, and it is private, discrete from that of every other individual. We can only observe first hand our own conscious processes. Our ability to interact with the environment depends on the accuracy of our internal model of it, and our ability to communicate with other individuals rests on the common features of our respective conscious models of the world.

If direct observation of the conscious processes of others is impossible, observation of our own is surprisingly simple. Consciousness is a unitary process. Stimuli reaching different structures of an individual are easily and effortlessly related in consciousness. For example, unfamiliar shapes projected on one side of the eye can be recognized on the other, even though different sets of neural fibers leading to opposite areas of the brain are involved. As another example, shapes can be felt and then recognized by sight. Mind transcends particular bodily structures. It cannot be explained in terms of the activity of restricted parts of the organism's nervous system, for it integrates physically discrete events which occur in separate parts of the body whenever these events reach consciousness. For this reason, many present-day scientists conceive of mind as something akin to an energy field embracing the entire organism, and transcending its particular structures.

2. The Content of Consciousness

The content of conscious life is qualities: over-all qualities of feeling, like joy or fear or anger; specific sensory qualities like red or hard or salty or high-pitched; qualities of objects or people like beautiful or dangerous or honest. These qualities of experience are the material from which our internal model of the external world is made. As the physical world is a structure or organization of energy, the internal world is a structure or organization of qualities.

Qualities of experience do not exist in the physical world. They are the material of our model of the world, not of that world itself. However, we ascribe to the external world properties which are qualitative and refer to our internal model. The redness of a book on my desk appears to me a property of the book, yet I know that the redness I see is a sensory quality with no physical existence. True, the book is reflecting long wave lengths of light and absorbing others, but this physical fact is not the quality of experience "red". True, the

nerve fibers of the retina respond differentially to wave length, so there is a pattern of neural discharge which corresponds to light composed of longer as opposed to shorter wave lengths, but <u>this</u> physical fact is also not the quality of experience "red". The information this neural discharge contains is only used in building this red object in my internal model of the world. Were I color blind, or completely lacking color vision, an insect, or a creature from outer space, the book would appear very different. The red opaque book I see is a personal creation, part of my model of the external world, but not to be confused with that world. I am able to communicate with and understand other individuals only because their model of the world is similar to my own.

The qualities of conscious experience are the building material employed in the four major classes of conscious activity -- feeling, sensing, remembering and thinking -- that are to be discussed. Each of these contributes to the control process as it originates within the human operator.

3. <u>Feeling</u>

Those qualities of experience called the "basic emotions" or just "feelings" comprise the most primative conscious events, and underlie the whole of man's experience and behavior. We recognize them by their undifferentiated, often intense character. Feeling qualities are pervasive; they involve the whole organism. They provide the motive force, the energy behind the control process, and the means by which we evaluate its outcome. In the final analysis, man is moved to modify the environment by the way he feels about it, and he passes judgment on the modifications he makes in the same way.

4. <u>Sensing and Perceiving</u>

As living organisms evolved, the relation between the organism's feelings and certain physical events affecting the organism developed. Certain qualities of feeling became triggered only by specific patterns of energy or sequences of events. The presence of the given quality of feeling thus became a signal for the events. This process marked the beginning of the sensing function. It proved enormously useful to the evolving organism, and developed, as we know, in the most elaborate way. Specialized structures evolved that were responsive to particular aspects of the environment, and produced their own specialized conscious qualities, e.g., odors, tastes, sounds, colors, and shapes. These qualities carried information about the external world.

As the sensing function developed, sensory information assumed less the character of special signals and more that of a model of the environment. This marked a long and important chapter in evolution. The unitary nature of consciousness served always to integrate sense data from all modalities. A model of the environment created only from sensory qualities has one severe weakness, however: it is bound to present time. Sensing signals refer to external events of the moment. Perception and understanding of the environment could occur only with conscious processes free from this time restriction. This begins to occur in remembering.

5. Remembering

It is one of the most curious and extraordinary facts of conscious experience that it has reference not only to processes occurring in the present, but also to events of the extended past. Conscious experience makes use of present sensory processes and traces of past experience in building a model of the environment that is not tied to the present. We see a face: if it looks familiar, if we recognize it, if it reminds us of someone, or even if we only recognize it as a human face, the remembering as well as the sensing function is at work.

The past is gone. The fact that past events play such a major role in conscious life owes to some sort of structural residues of past experience which are utilized in remembering. The nature of memory traces has been the subject of widespread speculation, but little is really known about it. The physical basis of memory traces remains a major scientific mystery.

Sensing involves the creation of qualities of experience based on the neural response to forms and patterns of energy. In remembering, traces of past experience are energized and they, like patterns of sensory neural response, serve as the basis for the creation of qualities of experience. While energy patterns from the external world bring the organism information about present events, energized memory traces provide it information about the past. Remembering may be considered a process of sensing appropriate memory traces.

Given an organism that creates conscious qualities in response both to patterns of sensory-neural discharge and to memory traces of past experience, a major problem concerns the way in which appropriate traces are selected and energized. The amount of information accessible to memory is incredibly large. The means by which appropriate selection is made from this mass of information is perhaps

the major problem in understanding the process of remembering. Anyone who has ever struggled with problems of filing and retrieval in an office or library must have wondered at the remarkable speed and efficiency with which these functions are performed in human remembering. How is the contact between particular conscious events and relevant memory traces established?

The individual is probably never twice exposed to the same conditions of stimulation. For example, in terms of the pattern of energy striking the retina, a familiar person never looks the "same" twice, yet recognition takes place most often with remarkable ease. Remembering is not tied to the particular sensory patterns that give rise to it. Rather, the sensory patterns give rise to the creation in consciousness of a model of some aspect of the environment, and it is the model, which is largely independent of the original sensory patterns, which is remembered. When a similar model is created in the future (regardless of the stimulation giving rise to it) traces of the earlier model are aroused. The basis of the arousal is qualitative similarity between the percept (new model) and the traces of the past. Memory traces are laid down and aroused by means of this qualitative similarity principle. Certain Gestalt psychologists have suggested that trace arousal is accomplished by a kind of resonance, implying perhaps a characteristic frequency associated with qualities of experience that arouses traces laid down at similar frequency in the past. This is an intriguing speculation but hardly more at present.

Remembering frees conscious processes from present time, and allows the model of the environment to incorporate temporal as well as spatial relations. The simple relation of succession, "event A is always followed by event B" would be forever beyond the grasp of the organism which retained no trace of event A when event B occurred. Yet the grasp of such simple temporal relations marks the beginning of the understanding of the relation of cause and effect, and so of the control of the environment.

There are many types or levels of remembering, from the simple recognition of an object that is present to the senses to the recreation in consciousness of remembered objects or events. The latter process, which may be independent of present stimulation, is the most highly evolved. It closely resembles the still more developed process which we have labelled simply "thinking".

6. Thinking

We cannot add up feeling, sensing and remembering, and obtain something that looks like human conscious processes. The most characteristic feature of mental life is not included in these functions. The distinguishing feature of human thought is that it is not bound to either the experienced present or the remembered past, but is free of both. The qualities of experience involved in perception and memory are used in new combinations in thinking. The model of the world which is created in consciousness may include the near or the remote in time or space, the real or the imagined, the past, present, or future. Man creates models, not just of objects, but of actual and possible courses of happenings, of possible future states of the environment and the events necessary to bring them about.

The mental modelling process operates on a fast-time rather than a real-time basis, in that events may be thought about much faster than they occur. Events are schematized and compressed in consciousness, with significant points and end results included but much of the remainder omitted. The human operator in a control system may consider several possible courses of action in less time than it takes to carry out one of them.

B. Conscious Processes in Manual Control

The significance of the fully developed capacity to model real and imagined environments in consciousness goes far beyond those factors that are crucial for understanding manual control. The development of symbols, objects of thought which represent things other than themselves, is in itself one of the most significant chapters of human development. Vicarious learning, i. e. learning via symbols or other indirect experience is another. Manual control, involving as it does an immediate relation between an individual and some variable in the environment that is controlled by the individual, is concerned with certain aspects of the thinking process and not with others. If thinking is for our purposes defined as building in consciousness a dynamic model of real and imagined events in the environment, then manual control is especially concerned with these aspects of thinking:

1. Goal conception, i.e., predicting (envisioning) possible future states of a controlled variable, (a) if nothing is done to affect it, and (b) if certain of the available control actions are taken.

2. Goal selection, i.e., planning the desired future state of a controlled variable by choosing from the range of possibilities perceived. This selection is made with reference to appropriate criteria for making a choice.

3. Programming the hierarchical sequence of events in the environment required to bring the desired state about, together with the control actions required to initiate and carry forward this sequence.

4. Carrying out the programmed sequence.

These together comprise the activities of the human operator in a manual control task. Discussion of the third and fourth is reserved for the following chapter.

1. Prediction

It is commonplace in engineering literature to state that control systems function so as to reduce the difference between their input and output, i.e., the desired and actual values of the controlled variable. This is not true of the human control process nor of man-operated control systems. Because of the operator's ability to think, to extrapolate forward in time, the manual control process is oriented around the future. This can be stated as a basic characteristic of manual control:

> Manual control systems function to reduce the difference between what an operator wants to happen to a controlled variable and what he thinks is going to happen unless he institutes a change.

What the operator wants to happen reflects the operator's planning of the desired future state, as discussed below. What the operator thinks is going to happen represents the goal conception or prediction process, the extrapolation into the future of the state of the controlled variable in the operator's dynamic internal model of the environment.

Prediction thus plays a key role in the process of manual control. The freedom of conscious processes from present time, the ability to envision the future, forms a major part of manual control skill. Since the operator controls a complex system by reducing the discrepancy between what he wants the system output to be and what he predicts it

will be if he institutes no change, his ability to predict is crucial to the success of his control.

The primary difference between the experienced operator of a complex control system and the novice is in the ability to predict. When a novice is given displays that enable him to predict system output accurately, he performs like a skilled operator without the need for training.[1] The principal reason for the long learning time needed to train pilots, helmsmen, and other operators of difficult control systems is the time it takes to learn to predict system performance under various conditions. The ability to predict system performance is in major respects the same as the ability to control the system.

The automatic controller operating only in real time and lacking the capacity to predict is likely to require the correct adjustment of multiple feedback signals to control a complex system. The combination of these signals in the controller provides what is called the "compensation" needed to bring about satisfactory control in the face of the dynamic characteristics of the system. Some workers in the field of manual control have applied the terms "compensation" (or "equalization") to control by the human operator. However, when the operator is in a situation where he is able to predict system performance, he need not "compensate" or "equalize"; the ability to predict makes these unnecessary. By moving his control so that predicted system performance corresponds to what is desired, the operator controls without "compensation". Only when an operator is performing in a situation in which prediction is not possible can the concept of "compensation" be applied with validity to manual control.

It is tempting to the control engineer to say that the ability to predict makes it possible for the human operator to "compensate". He can demonstrate mathematically that predictive information provides the information for compensation that is required by the operator. This is an inversion of what I believe to be the correct view. Control began with man and was extended to automatic devices; we should not explain away the basic characteristics of the control process in man by the less basic concepts developed to describe automatic control processes. The concept of prediction is more basic in control than the concept of

---

[1] Kelley, C. R. Developing and Testing the Effectiveness of the "Predictor Instrument", Office of Naval Research Technical Report 252-60-1. Stamford, Conn.: Dunlap and Associates, March 1960.

compensation, which is appropriate for a device that does not predict. I believe the proper viewpoint is that in a control device limited to combining present time signals, i.e., unable to predict, the device must compensate for dynamic characteristics of the system for stable control to result. This compensation is a substitute for prediction, rather than prediction being a form of compensation.

Predictions of more than one kind play a role in manual control. Control action taken by the human operator is shaped by the difference between the desired state of the controlled variable and that predicted if no action is taken. Also involved is a prediction as to what will happen if a particular action is taken, i.e., "if I take this action, then I predict this will happen to the controlled variable". Control thus employs not only a prediction of the result of no control action but of various possible control actions. In particular, the operator plans or programs an action or sequence of actions in control based on a prediction as to what the result will be.

The selection by the operator of the desired value of the controlled variable, the operator's plan, is limited by the prediction process. The operator plans system output by selecting, from the possibilities he perceives, that which he considers desirable. What he perceives as possible depends on the prediction process. In continuous control, what is possible may often be defined by what would happen if the operator moved a control to either extreme of a range of operation. These extremes define the range of possible system outputs. The operator selects his desired output or plans his course within what he believes (predicts) the possible range is.

Figure IV-1 illustrates the role of the prediction process in steering an automobile. The dotted line represents what the operator expects would happen if he did not move the steering wheel. Because this would bring him away from his desired course, he is instituting a correction, which he predicts will bring him to the path centering on the dashed line. The two diverging solid lines represent what he expects would happen were he to turn full left or full right. They define the limits within which he plans his path.

The dashed area represents the desired path within which the driver wishes his car to remain while going around the curve. This is the driver's plan. It is simple here, in the absence of traffic, but can become complex. It is differences between the desired path and what the operator predicts will occur if he institutes no change (dotted line) that are together most useful to the driver in programming his response.

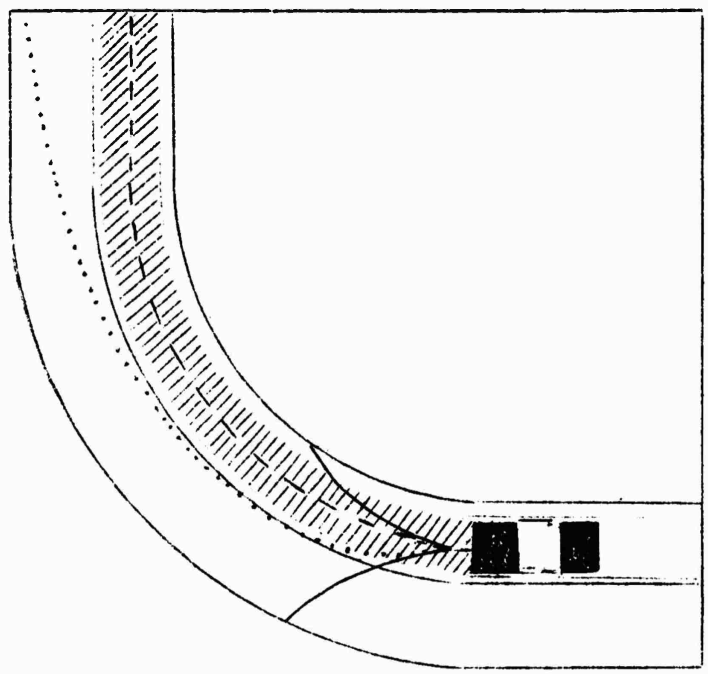


Figure IV-1. Prediction in steering an automobile.

2. Planning (Goal Selection)

Planning in the control process is the conception of and choice among possible future states. As such it is the central activity in manual control. The ability to predict is a necessary condition for planning, defining what future states are possible, and what must be done to realize them. The process of planning is the charting of a course within the limits of what is perceived as possible.

Planning may take place on each level of a hierarchical control process. Consider the example of the previous chapter in which the environmental variable X is controlled by means of Y which is controlled by Z, Z being the variable directly affected by a control system. In a completely manual system the operator must plan not only the outer loop variable X, but also the intervening inner loop variables Y and Z by means of which X is controlled. There are thus levels of planning (and levels of prediction) for each loop in the manual control process.

The output of the human operator, the bodily movements activating the control process, also require a form of planning. This innermost loop of the manual control system differs perhaps more in degree than in kind from the longer range planning in the outer loops. Since it does serve as the transition point between the conscious processes of planning and the fundamental human output, bodily movement, it is termed "response programming" rather than "planning", and is dealt with in the following chapter.

Planning as here defined is focussed on the environment, and the changes to be made in it. It is first and last a form of thinking, and is the root of all control processes, manual and automatic. It presupposes an awareness of the environment, a desire to change that environment in particular ways, an ability to foresee possible changes and their consequences, and the motivation to do what is required to put a selected change into effect.

It is generally true that thought tends to become focussed on the higher outer levels of control as a manual control process is learned. This is only to say that as the control process becomes more practised, the operator can devote more attention to long range goals, to ends rather than means to those ends. Shorter range lower level inner loop processes tend to become habitual. This rule is true only within definite limits, however. It cannot be generalized that outer loop activities are necessarily the primary focus of attention in highly practised manual control activities. When the goal of the outer loop

process has been chosen, it may govern inner loop processes without remaining the focus of attention, for attention often must be directed to the means for achieving the goal. Conscious processes are, in fact, invariably concentrated on whatever levels of control at which choices are being made, be these in inner or outer loops. When the outer loop goal is fixed and is no longer to be chosen, attention can move to the inner loops, where the route to reach the goal is to be chosen. When inner loop activity, in the form of highly practised movements of a skilled operator, has once been learned, the responses become habitual and in a sense automated, and the operator's attention can move to more outward loops, to consider the goals towards which his inner loop activity is carrying him. When no choices are being made at any level, the entire manual control process may proceed habitually or automatically at all levels with little requirement for conscious attention.

Thus, at whatever level there is novelty in the control process, i.e., unexpected events, choices for which there is no habitual response, conscious processes are brought into play. Consciousness bears a clear relation to choice. Only where there are alternatives to choose from has consciousness a function in the control process.

3. Criteria for Selecting a Plan

Planning in manual control involves the choice of alternatives, or the choice within a range of alternatives. Prediction provides information as to possible alternatives, but this type of predictive information alone is not enough to guide the choice that is made. In addition, there must be criteria for judging the anticipated consequences of the alternatives available.

A criterion for choosing an alternative in manual control implies a goal apart from or at a higher level in the control hierarchy than the output of the control system. This new goal, however, is achieved through the same control process. Furthermore, it imposes requirements on that process. Not only is the control system controlling the environmental variable X, it is doing so in such a way as to minimize, maximize, or optimize some other variable A, where A is the criterion. Frequently there is more than one criterion against which an operator evaluates alternative paths open to him in control, and certain of these criteria defy analytic treatment. It is not always possible to know why one alternative was selected over others. An operator may select an alternative in control because it appears easier, perhaps because it requires only a highly practised response. On the other hand, he may

choose an alternative that is novel because it appears more instructive, interesting, or enjoyable. In a dangerous maneuver, human safety is a major criterion. Reliability of equipment under various possible operating conditions is another criterion that is at times of pre-eminent importance.

The most discussed criteria and those most amenable to study relate to system performance. An alternative might be chosen in vehicular control because it promises quicker arrival or less distance travelled. In general, one control activity may give promise of more precision in control of the environmental variable than others. In systems where fuel expenditure is critical, the choice which looks as if it will use less fuel may receive priority over all others. System performance criteria usually have the advantage that they can be measured and evaluated quantitatively. Control systems which minimize fuel, mean square error of output, probability of exceeding equipment limits, or similar kinds of criterion functions can be developed employing quantitative methods. This more than any other factor accounts for the relative emphasis on this type of criteria in control.

A criterion has been earlier defined as a higher level goal against which the control process is evaluated, and for there to be a criterion there must be a desire -- or a necessity -- that this goal be reached. The most urgent form of criterion is that which is both necessary if the control process is to succeed and yet is difficult to meet. Such a criterion can become an overriding consideratoon in a control process, and the design and operation of a control system can be oriented around this one variable. The great majority of possible control activities may be eliminated from consideration because they fail to meet the given criterion. The range of choice in control thus becomes greatly restricted.

A concrete case in point is the fuel consumption criterion for spacecraft maneuvers. Spacecraft are vehicles which operate with severe limitations on the available fuel. This one consideration assumes overriding importance in their design and use. The primary limiting factor on the space missions attempted is that of fuel. This country cannot yet perform missions it would like to perform, e.g., send a team of scientists to circumnavigate Jupiter or land on Mars, principally because of our limited ability to lift vehicles carrying the necessary amounts of fuel off the earth's surface and into space. Every other problem in spaceflight has been dwarfed by this one. The effects of the limitation are pervasive, and are easy to overlook. Why is the reentry of spacecraft into the earth's atmosphere from a lunar

or planetary mission such a delicate maneuver? -- Because earth's gravity imposes such high entry speeds on vehicles coming in from a superorbital trajectory, and we cannot send enough fuel along to slow them down. Why must a spacecraft such as the Apollo lunar vehicle carry a highly complex guidance computer and navigation equipment of extreme precision? -- Because the fuel carried permits so little error in the Apollo trajectory. A twenty-first century astronaut with a controlled fusion source of energy in his space vehicle will be able to fly to the moon and back without need for such precision of navigation or the worry about reentry heating.

This example illustrates how a criterion based on a critical and difficult system requirement can become a predominant variable in a particular control process or system. It also illustrates how secondary criteria may develop from a more fundamental one. The accuracies required in space vehicle reentry and mid-course navigation become established and function as criteria in their own right, although they are derived from the more fundamental requirement that the mission be performed with severely limited fuel.

Criterion variables add an additional factor to the prediction requirements of manual control. Not only must possible outcomes be predicted in terms of the controlled variable and the intervening variables that affect it, but the way in which these possibilities satisfy the criterion must be predicted as well. This can sometimes be more difficult than predicting the control process. It is often hard for an operator to know what safety, reliability, RMS error, etc. he can anticipate with the various output possibilities at his disposal.

In the usual case in manual control there is not one but many criterion variables, no one of which predominates to the exclusion of the others. The spacecraft example is the exception rather than the rule. Manual control activities are apt to involve time and accuracy and economy and safety and reliability, and even the interest, ease, and enjoyment of the human operator. The operator is required to satisfy all his various criteria within reasonable limits, but a wide range of latitude commonly exists. The balancing of one criterion variable against another in the manual control process can be one of the major functions involved in human control operations. The best operator is not the one with the highest degree of manual skill, but the one that can be relied on to make the best choice of alternatives in control when complex criteria must be weighed and balanced against each other. The best operator is the one who can choose alternatives in control, can plan, taking into account the whole spectrum of criteria -- time, accuracy, safety, economy, and the others -- balance them appropriately and make the right choice. This choice is the plan or the desired path around which the human control process is oriented.

V. THE HUMAN OPERATOR: 2. MOTOR PERFORMANCE

> The creature does not merely move in a certain direction, like an inert mass impelled by an external force; its movements are quite incapable of being described in the language with which we describe mechanical movements; we can only describe them by saying that the creature strives persistently towards an end ... which end is generally some change in its relation to surrounding objects .....

William McDougall
Social Psychology
4th Edition, p. 354

When the human operator has selected his plan, which embraces the desired changes in the controlled variable and the sequence of events or processes that can bring it about, when he has examined possible future states and, applying his criteria, made his choice among these states, there remains for him to carry the plan out, to put it into effect. This involves human muscular activity, the sequences of bodily movements which initiate and correct the control process. These sequences are programmed in advance on the basis of the desired plan. The plan is then executed by means of bodily movements. The bodily movements involved in manual control are often highly skilled. The subject of motor performance is therefore considered under three headings:

A. Response Programming

B. Human Movement

C. Manual Skill

A. Response Programming

The final stage of planning in the human control process is the "planning" of the bodily activity required to bring about the desired changes in the environment. This is the innermost loop of the manual

control process. The output of this stage orders the human movements which initiate and regulate manual control processes. As the last stage in the planning process, response programming differs in important respects from planning activities which precede it. Its position in the sequence of control processes means it takes place after possible modifications in the environment have been considered, after criteria for choosing these modifications have been applied, after the desired values of the controlled variable and all the intervening variables between human output and the controlled variable have been chosen. There remains to be "planned" the bodily activity to bring the desired values of these variables about. This "planning" consists in large part of setting up a movement pattern or sequence of patterns associated with the desired changes. This sequence forms the program of operator activity. The bodily movements in the pattern may be triggered by observation of changes in the controlled variable. The process of setting up the sequence of body movements appropriate to achieve the planned changes in the controlled variable is what is here termed "response programming".

1. Complexity of Manual Skills

Simple movements that have been learned in the normal process of development such as reaching, grasping, pressing, twisting, etc. can be programmed in advance with no practice. The operator is able to originate a time sequence of a few such movements appropriate to achieve the planned changes in variables he controls. The movement sequences in the response program may in this case be nonhabitual. The program is developed as the need arises.

More frequently the movements involved in manual control are more complex. Highly skilled heavily practised complex patterns of movement play a major role in most manual systems. The operator may have built up a large repertory of skilled movement patterns or sequences which can be evoked when needed. In response programming involving complex skills, the program may incorporate such patterns without the necessity for planning the movements that make up each one. If we think of the response program as being somewhat analagous to a computer program that is being continuously developed by the operator, then the complex prelearned patterns of response in the operator's repertory are like subroutines. A subroutine is a pre-existing functional sequence of instructions stored in a computer's memory that, e.g., take a square root or perform some other common operation. A programmer can make use of subroutines available to him without the necessity for generating anew the sequence of instructions it involves.

In the same way, the control system operator can make use of prelearned patterns of response without the necessity for programming the movements making up the pattern. Like the computer subroutine also, the prelearned pattern of response is likely to be efficient, much more so than if the operator did have to originate the movement sequence.

As an example, consider shifting gears manually in an automobile. This is typical of complex movement patterns that occur in many control tasks. It involves coordination between one hand and both feet, while at the same time the operator must steer with his other hand as he watches the road ahead. Having learned this complex and difficult pattern of movements, the driver has it available as part of his repertory of learned patterns of response, and it can be inserted into the operator's response program with no need for separate planning of the movements that make it up.

The output of the planning system establishes requirements that the body must meet if control is to be achieved as planned; the response programming function develops the planned sequence of bodily activities to meet these requirements. As an inner loop activity, the body's response tends to be more quickly changing or higher in frequency than the outer loop changes it brings about. A single example from driving an automobile will suffice to show this relationship:

Outer Loop Plan	Response Program
Stop at signal ahead, then turn right	Move steering wheel as required to steer in lane
	Move foot from accelerator pedal to brake
	Depress brake with force needed to stop at intersection
	Depress clutch pedal before complete stop
	Hold brake and clutch pedals down
	Move shift lever from third to first gear
	Lift turn signal indicator lever to signal right turn
	Move right foot from brake to accelerator

Turn steering wheel to right

Release clutch gradually with left foot, and

Depress accelerator with right foot

Allow steering wheel to return to center

Move steering wheel as required to steer in lane

Let up accelerator

Depress clutch

Move shift lever from first to second, etc.

This example illustrates something of the nature and extent of the bodily activities required for what is, at the outer loop level, a simple and straightforward plan. Functional activities, such as "move shift lever from third to first gear", can each be broken down further into fine details of muscular contraction and release. Response planning never reaches to this ultimate level of bodily response. To return to our analogy between response planning and digital computer programming, the instructions in the computer (e.g., add the number in location 688 to the number in the accumulator; store the number in the accumulator in location 972) correspond to simple movements such as are employed in programming the organism's response. The particular muscles which must contract and release in coordinated fashion to execute these simple movements correspond to the detailed inner working of the computer which carry out each instruction.

2. Consciousness and Response Programming

When the operator is responding with prelearned sequences of movements, the individual movements of the sequences follow each other without conscious direction, and usually with little direct awareness of the movements in progress. The small degree of consciousness associated with skilled movements reflects the fact that there is little choice taking place as one movement succeeds another in the sequence "automatically". Consciousness occurs where there is choice, and in the highly learned movement sequence, the choice occurs in selecting and triggering the sequence itself, rather than in executing

the movements that make it up. Conscious attention can, in fact, degrade or disorganize the learned pattern of response. It is only when a novel response pattern is called for, where no appropriate ready-made response exists, that response programming takes on the characteristics of the conscious planning that is common in the more outward loops in the hierarchy of control. These exceptional cases when a novel response pattern is called for may be highly important, however, e.g., the emergency situation.

The fact that learned patterns of movements are carried out without conscious direction in an "automatic" fashion should not be taken to mean that the sequence is performed blindly and mechanically. The habitual sequences of movements are tailored and modified to fit the particular situation in which they are used. The highly skilled driver has available to him the complex patterns of movements connected with the many subtasks of driving: shifting, braking and accelerating, patterns of steering as in turns, passing, etc. Yet it is true that the "same" movement pattern is never repeated in precisely the same way twice. Turns are made more or less sharply, braking is adapted to the exigencies of traffic, gears are changed at approximately the same speeds on the same slopes, but not precisely so. Even with highly practised movement sequences there is some modification of the sequence with the situation, some element of choice and of short range planning; and to just this limited extent there is some degree of consciousness associated with the movement pattern. The driver proceeding "automatically" has some level of awareness of the environment, and does shape his movement patterns to fit it.

3. Prediction and Response Programming

The response program as it is formed by the operator becomes available in the output prediction process. The prediction that can be made takes the form, "if I follow the response program, then this is what I expect to happen to the variables I'm controlling". This is the last and most highly evolved prediction the operator can make, involving as it may a complex sequence of control actions. It is a relatively short-term prediction, concerned only with the period ahead for which movements are already planned. In typical manual control tasks this may amount to a few seconds. Discrepencies between what is desired and what is predicted can then result in final revisions or adjustments to the response program. Discrepencies should usually be small at this stage, because the response program was itself based on the desired values of output variables. This constitutes a vernier adjustment in control, a fining down of control that had been previously planned more grossly.

4. The Convergent Nature of Planned Responses

The "fining down" of planned responses as the time for their execution approaches is an important feature of the process of control. The predictions on which control is based are inaccurate as a function of length of the span of prediction. Outer loop goals of control are formulated first. At the outset there may be no more than a vague premonition of the actual bodily activities that will be required to reach the goal. As more inward loops define subgoals and paths to reach them, the bodily movements to be required become more closely specified. By the time the innermost loop is reached, a precise complex programmed sequence of bodily movements is defined. This sequence brings about subgoal after subgoal until the goal is reached. Whereas the outer loop goal may be explicitly defined minutes or hours or more in the future, the inner loop program is formulated only vaguely this far in advance, and may not begin to become definite until a few seconds before it is scheduled to occur. In these few seconds, adjustments and corrections are made to tailor the movements to the situation. The increasingly precise response program makes possible a more accurate prediction of the outcome when the program is carried out. Small discrepencies between predicted and desired outcomes can still be eliminated by minor modifications in the movement pattern. The poorly defined sequence of activities planned at the outer loop level thus turns into a detailed, precisely adjusted programmed sequence of body movements that is appropriate to achieve the long range goal when the inner loop programming activity is completed.

B. Human Movement

Every human control activity is traceable to human movements. The inner loop of planning activity is "response programming", the units of which are movements or patterns of movements of the body. In manual control, movements of the limbs are of primary significance. The body moves as a consequence of coordinated patterns of contraction and release of hundreds of muscles, composed of hundreds of thousands of muscle fibers. These coordinated patterns of contraction and release are governed by underlying patterns of neural response, which reach muscle fibers via the nerve fibers or cells, known as neurons. Body movements are thus directly traceable to patterns of efferent neural activity that must be modified to reflect the "response program" of control movements. Among the least understood aspects of the manual control process is the means by which a program of responses, which is a scheduled or planned sequence of movements, brings about the required pattern of neural activity.

The organism is a unit, a functioning whole, which is differentiated into its incredibly detailed and intricate structure without losing its unitary nature. Just as the organism perceives and thinks as a unit despite the separate structures that mediate its awareness of the environment, so it moves as a unit despite the hundreds of millions of cells that must coordinate in such movements. The true unit of life is not the cell but the organism.

Higher organisms are composed of billions of cells of many types, each type having specialized functions. Muscle cells are specialized to contract, nerve cells to store information and to communicate, i.e., transmit impulses from one part of the body to another. These together form the basis of movement of and within higher organisms.

1. Muscles

Muscles are made up bundles of long thin fibers, individual fibers ranging from a small fraction of an inch to more than an inch in length, and on the order of 1/10,000 of an inch in diameter. The fibers are embedded in connective tissue that provides for their maintainance via blood circulation, and through which the nerve supply to each fiber passes.

Muscle fibers are specialized to contract, and when the attaching nerve ending stimulates them to do so, they become shorter and thicker, a change occasioned by an actual rearrangement of the atoms within the molecules of the fiber. The muscle fiber's contraction is highly energetic in comparison with the energy of the nervous impulse triggering it, consuming millions of times as much of the body's energy. The muscle fiber has its own supply of metabolic energy, of course, and receives only the impulse to contract from the nerve cell. The connection between nerve and muscle is perhaps nature's original control junction. The exact nature of the triggering action is not entirely certain, but it may be either electrical or chemical.

2. Nerves

Like the rest of the body, nerves are composed of myriads of discrete cells. Nerve cells are called "neurons". These may be longer and thinner than even the individual muscle fibers. A single motor neuron may extend from within the spinal column into the foot, and be no more than a few microns in diameter. The terminal end of the motor neuron, the axon, splays to attach to muscle fibers. The receiving end of the neuron is usually branched into a number of

"dendrites" which receive stimuli from the axons of many neurons leading to them. The axon-to-dendrite connecting links between neurons are known as "synapses".

Neurons are specialized to conduct. Impulses reaching the neuron via its dendrites may trigger the action of the neuron. Usually at least two impulses seem to be required. Once triggered, the neuron responds by transmitting an impulse that is independent of the strength of the impulse received. This is the well-known "all-or-none" principle of neural transmission. The analogy is to firing a bullet: the explosion produced is independent of how hard the trigger was pulled.

It is tempting but misleading to think of neural transmission as an electrical phenomenon. Electricity travels with a velocity several million times as fast as neural transmission. Impulses move through neurons at velocities ranging from the speed of walking to that of a slow aircraft, e.g., from four or five up to about two hundred miles per hour. The impulse consists of what has been called a "propogated disturbance", and is an electrochemical chain reaction that moves along the nerve fibers.

There are literally billions of individual nerve cells in the body, organized in the most complex way and in continual activity of one sort or another, and there are a range of control activities operating within the nervous system itself. There are reverberating circuits in the nervous system, capable of perpetuating themselves indefinitely; there are neural "signals" which can inhibit or suppress as well as enhance a given neural activity. Neural inhibition plays an important role in coordinated movement. The picture here should be of hierarchical control processes within the nervous system itself, the lower levels of which carry on continuous activity, modulated by higher level input signals that enhance or suppress the level of activity as appropriate for carrying out the higher level process.

3. The Neural Basis of Movement

The axon or terminal end of each individual motor neuron splays out to supply connections to a number of individual muscle fibers, how many reflecting the grossness or fineness of the coordinations in which the muscle is involved. In the large muscles of the body there might be more than a hundred muscle fibers fired by a single neuron; in the fingers there might be only three or four. Thus the contraction occasioned by a single neuron is much larger in the case of gross movements than fine, and the degree of neural involvement is much greater for fine movements than gross.

The neural apparatus supplying the muscles provides the appropriate coordinated volleys of impulses to the muscle fibers which result in all of the body's activities, from simplest to most complex. The neural apparatus, like the control process it mediates, is organized hierarchically. The basic hierarchy of nervous organization is not built around individual muscles or limbs, but around patterns of body movements. Even the lowest level of the hierarchy, the body reflexes that involve but a single segment of the spinal chord, show this form of organization. Morgan and Stellar state:

> A segmental reflex is the simplest kind of behavior to be seen in any organism. Yet it is a highly organized affair. It is a pattern of response that is made up of many neuromuscular units, all acting together with precision of timing and amount of contraction. The pattern is not simply a mechanical contraction of a particular group of fibers but rather a pattern of movement.[1]

The nervous system, then, is developed about patterns of movement of increasing levels of complexity. Electrical stimulation of the pyramidal section of the cortex of the brain evokes much more complex patterns of body movements than segmental reflexes. The different movements evoked by stimulation of different points in the brain involve many of the same muscles, and thus stimulation of a point in the brain arouses patterns of movement of different muscles, and the same muscle can be stimulated from different points in the brain.

4. The Mystery of Neural Functioning

We know a great deal about the intricacies of neural functioning, but the most fundamental questions remain unanswered. As the previous chapter indicated, patterns of sensory neural response form the primary input channel to the dynamic model of the environment which is our perception and understanding of the world around us. How the translation from coordinated patterned volleys of sensory neurons to our perception of the world takes place, we do not know. Similarly, our perception and understanding of the world makes it possible to modify it. We have also described processes leading from our dynamic

---

[1] Morgan, C. T., & Stellar, E. Physiological Psychology. New York: McGraw Hill, Second Edition, 1950.

model of the environment to the organization of the response program consisting of patterns of movement to change the environment. This response program may be set up in some form of dynamic neural storage, ready to be keyed by perception of the appropriate situation in the environment. What this storage is, or how the program emerges to become the coordinated patterns of nerve impulses that we know underlie human movements is, like the forming of the internal model of the world, a mystery to us.

Thus, both the forming in consciousness of the internal model of the environment upon which the process of control is based and the patterns of bodily movements aimed to change this environment are not understood, although we know that both are intimately related to the coordinated patterns of neural discharge which underlie both input to and output of the individual. Both appear to this author as aspects of a single process, a creative activity on the part of the living organism. Perception is not a process of reception of stimuli, but an active model-building process directed outward toward the environment; the very process of building the model of the environment brings into existence the capability for changing it. In time science will come to understand the basis of this natural process. At present it does not.

5. The Hierarchy of Movements

The control process has been described as hierarchical in nature. It is not surprising then that bodily movement, on which the control process is based, should reflect this same basic structure. Human motor performance is organized about patterns of movement, as we have seen. These patterns, however, are of different levels, ranging from the simplest reflex to the most complex coordinated skill.

Many of the body's movement patterns are built into the organism at a level as fundamental as the body's inherited structures themselves, and it is from inherited structures that these patterns are no doubt derived. This includes not only the internal movements of the heart and other organs, but many reflex activities, including those by means of which we maintain our body position and balance in the face of gravity. Many of the movement elements of locomotion and manipulation appear likewise to be "built in" at the reflex level. It is difficult to separate the influence of learning from that of maturation in man, but it has been experimentally established that many complex movement patterns (such as pecking in chickens and swimming in salamander tadpoles) which appear to be learned are, in fact, inherited. The patterns appear fully developed in experimental

individuals deprived from birth of all opportunity to learn them. They appear only after necessary growth has taken place, but are unlearned. Many physiologists believe that complex patterns of human movement such as those of eye fixation, reaching and grasping, and even walking, are based largely on inherited structures. Because of the long period of immaturity of the human child, during which so much is learned, we don't know what man's basic inherited repertory of movement patterns is. In terms of function, however, it matters very little whether particular patterns of movement are inherited or ingrained by long habit. The organism does have available a repertory of short range movement patterns that include postural orientation, reaching, grasping, pulling, pushing, turning, pressing, etc., which form the inner loop of the hierarchy of human movement.

The next higher level of movement patterns involve the highly practised habitual skills of locomotion and manipulation. Everyone develops such patterns in learning to move about, feed themselves, and perform simple manipulations on the environment. These patterns include in man the use of external objects which the child begins to employ as tools or implements. Consider the types of skills being developed by the infant in the sand box playing with sticks, shovels, and pails. This level of movement patterns incorporates reflex and learned responses into coordinated activities which show increasing dexterity.

The adult takes it for granted that he can insert a key in a lock and turn it, screw the top on a jar, adjust the knobs and levers controlling a car heater, dial a telephone number, and all of the thousands of other patterns of movement that we perform each day with hardly a thought. Each of these activities is composed of chains of lower level reflexes and simple habits, organized into patterns which seem to us the simplest kind of movements.

More complex movement patterns include the special highly practised and highly habitual skills such as writing or playing music, and the more general flexible longer range patterns of movement made up of combinations and sequences of simple movement patterns organized around some goal. Consider getting into and starting an automobile as an example. This is composed of many discrete acts of body, hands, and feet, some of which require close coordination of different subpatterns of movement, i.e., we must push in the clutch before turning on the starter switch, and coordinate the activity of gear shift lever, clutch and accelerator to begin moving. These kinds of activities, three levels "up" in our hierarchy of movements, can, if highly practised, be incorporated into our response program with hardly

a thought. Movement patterns at this relatively high level are chained together to form the still higher level of planned activity sequences that make up human control activities.

To summarize the hierarchies of human movement, the specification of which is of necessity somewhat arbitrary, the following levels may be distinguished:

a. Reflexes and simple habits; postural reflexes, fixating, reaching.

b. Basic goal-oriented movements of locomotion and manipulation, such as walking, picking up something, pressing and turning.

c. Coordinated brief patterns of skill usually oriented around short range accomplishments, such as starting a car, drilling a hole, putting on a shoe.

d. Sequences of activities involving a series of short-range accomplishments organized around a longer range goal, such as going to the store, repairing a lamp, getting dressed for a party.

These differences in level correspond to the description of the hierarchies of control of Chapter III. The higher levels refer to longer range goals; the lower levels are shorter range, more quickly changing, higher in frequency, much more detailed. Human skill develops from the lower levels up, or from the inner loops outward.

C. Skilled Movement

"Skill" is a term used to refer to certain difficult learned patterns of motor (more accurately, perceptual-motor) performance. It involves precision and timing of movements that are oriented around a task or goal. Skilled movements are highly practised movements, carried on without excessive expenditure of energy. The highly skilled performer will do easily and gracefully what the less skilled will do awkwardly and with difficulty, and the end result, the goal, will be achieved as well or better by the more skilled.

1. Consciousness and Skill

Highly developed motor skills are performed with little conscious direction or planning. Before the skills were acquired, the same tasks

required a great deal of conscious attention. The skilled person thinks very little if at all about the motor aspects of the task. Thus the student driver thinks ahead, plans his movements carefully, and is very attentive to what he will do next -- and drives badly; the skilled driver maneuvers through difficult traffic without giving it a thought.

Another way of expressing this difference is in terms of the hierarchy of control. The novice must orient his thinking around as yet unlearned inner loop activities. He concentrates on these to the relative neglect of the longer range outer loop functions. In bicycling, the neophyte may have so much trouble maintaining balance that he can't pay attention to where the bicycle is going. The cadet flier may work so hard at keeping pitch and heading under control that he fails to stay at the desired altitude. However, as skill develops, the inner loop functions become habitual, and attention is freed for higher level functions. The response program for the task that is learned is built up and stored in some way, and is "there" at the right time. The sequence has become like the computer's stored subroutine, available whenever needed, but with one major difference: the response program governing the skilled movement is more flexible. It is tuned to the environment, and varies appropriately with it, so that the patterns of skilled movement fit situations and sequences of events that may never be twice exactly the same.

2. Role of the Senses in Skill

Patterns of movements in the human operator's response program are triggered and governed by patterns of sensory information. This is as true of the lowest level spinal reflex as of the highest planned movement sequence. Patterns of movements are keyed to the sequence of events within and outside the organism, are initiated, halted and modulated or modified by such events.

The human senses play two entirely different functions in human movement. The first is to bring in the information which makes possible perception, prediction and planning. This information appears in the input to each level of the control process. In addition to this input information there is the current feedback from the body and from the environment, provision of which is the second function of the senses. Feedback sensing information refers to the present situation. It is used to compare and confirm, to adjust and correct the plan or program of responses, and to time and trigger movements making up the response program. Input sensing and feedback sensing are best contrasted by noting that input sensing contributes to the development of

the operator's plan or program of action, while feedback sensing contributes to its execution.

Feedback information referring to longer range outer loop aspects of the control process affects the planning or programming process for more inward loops. This can be seen by inspecting a diagram of a hierarchical closed loop system such as that of Figure III-2 (p. 27). Feedback information is feedback on one level only, for it is incorporated into planning and programming of more inward loops. Thus the ship helmsman receives feedback information about his actual heading from his compass, which he compares to an expectancy, the latter incorporating sensing information from higher levels of the control process, e.g., navigation sightings. The feedback information about actual heading is compared to input information about predicted and planned heading, and modifies the planning-prediction process for more inward control loops, e.g., rudder angle control.

The planning-programming process builds an expectancy as to what feedback information will be received. Feedback information reflects departures from expectancy. It makes it possible to adjust and synchronize responses as needed because of variations of events from expectancy, and is valuable to the extent that there are such variations. If there were no variation from expectancy, the feedback loop would carry no information in the quantitative meaning of the term.

As skill increases, human operations are performed more and more uniformly. There is, in fact, a steady decrease in the range of variation of feedback with learning. It becomes possible to carry through sequences of skilled movements with much less reference to feedback information when the movements can be performed in a uniform fashion. Pre-programmed responses are carried out with much less reference to present events. As Bowen states:

> While the closed loop character of skillful behavior is its basic property, there is a tendency for behavior to develop open loop characteristics as practice leads to perfection. As the required actions become increasingly perfected, there is less need for feedback information ... the checking and corrective feedback is in operation only intermittently and partially.[1]

---

[1] Bowen, H. M. Human skills as systems considerations. In Javitz, A. E. (Ed.) Engineering psychology and human factors in design. Electro-Technology, May 1961, 29, p. 125.

This accounts for the fact that the human operator sometimes operates for brief periods in "open loop" fashion on very highly practised motor skills. The response program is so well learned that a sequence of movements can be carried out with little or no sensory feedback information.

The changes in the use of input sensing information that occur with the growth of skill are those associated with the decreased attention to inner loop and increased attention to outer loop activities. As conscious planning of the body's activities becomes unnecessary, attention to the environmental variables under control increases. At first the variables most immediately affected by the body's activity are the center of attention. With increasing skill the response program extends to include these more immediate results of the body's control activity, and the longer range, higher order variables in the chain of control come more and more into the forefront. With this change, better ability to plan and predict the longer range outer loop variables is developed. With the growth of skill, attention is progressively freed from short range inner loop processes.

To illustrate, the operator learning to control depth of a submarine goes through a steady progression in time spent attending his various instruments. He first learns to control the angle of the "planes", the horizontal rudders by means of which depth control is achieved, by moving a handwheel regulating the rate of motion of the planes, and watching a plane angle indicator.[1] This is quickly acquired. Then plane angle is employed to control the pitch angle,[2] a higher order function. This is much more difficult, and requires considerable practice. Finally, pitch angle is used to control depth. As learning proceeds, the amount of time spent looking at the inner loop display (the plane angle indicator) gradually reduces, and that spent looking at the outer loop display (the depth gage) gradually increases. At the same time the operator's ability to predict submarine depth is greatly increased.

---

[1] Many modern submarines have position control rather than rate control of the planes, i.e., the displacement of a control stick or wheel brings about a proportional displacement of the planes, rather than a proportional rate of their motion.

[2] Submariners refer to the dive or rise angle of the submarine as "boat angle" or "trim angle". Pitch angle is the term preferred in the engineering literature.

3. Other Characteristics of Skill

Skills have been studied from the earliest days of psychology, and there is a substantial literature dealing with the subject. Some of these characteristics or properties of skill should be at least mentioned here, though no attempt will be made to cover them all.

a. Fixed vs. Ballistic Movements. A fixed movement is made with tension in antagonistic or opposing muscles, movement being made by continuous adjustments in the relative degree of tension. Writing, for example, involves this type of motion. Ballistic motions involve contraction of one set of muscles and relaxation of its antagonists, as a result of which the movement (usually of a limb) takes the form of "throwing" the affected part of the body from one position to another. Swinging a golfclub or a baton illustrates the ballistic form of movement. The distinction between the two is not always clearly drawn, and most complex patterns of movement involve both. As skill increases, however, there is a general tendency for the ballistic type of movement to increase at the expense of fixed movements. This tends to result in more graceful movements involving smaller expenditures of energy.

b. Rhythm. Skilled movements tend to be precisely timed, and repetitive skilled movements often develop a rhythmic quality. Rhythm helps the timing of movements, and thus increases their uniformity and precision. Patterns of skilled repetitive movements tend to develop a time structure that enables performance to be carried out by rhythmic ballistic movements. The skater exemplifies this quality in skilled movement. Physics supports the idea that rhythm contributes to the precision of movements as well as to their timing and succession. If a given ballistic movement is made to occur more uniformly in time by virtue of rhythmic performance, then force must be applied uniformly, for a larger force would cause a faster movement. The amount of force can thus be regulated by the rhythmic timing of the movement.

c. Reaction Time. Man reacts very slowly to stimuli when their occurrence or timing cannot be anticipated. Even a stimulus that is but a trigger to a single pre-set and ready response involves a delay of on the order of 0.2 seconds, and more complex or unexpected stimuli involve longer delays. If, however, the stimuli can be accurately anticipated, no delay need occur.

d. Committed Movements. There is a certain "point of no return", when a movement cannot be modified. Since movements are pre-shaped and scheduled, this is perhaps not surprising. There is a response

program in existence, and should it suddenly be replaced by another one, the process requires something more than a reaction time to initiate. While this change takes place the committed movement continues to run its course. Depending on the timing of the stimulus, the movement either is committed and runs its course, or does not occur. A part of the skill of boxing or fencing consists of feinting to get the adversary to commit himself to movements that will open up his defenses.

e. Discontinuities in Continuous Movement Patterns. Most complex skills involve various discontinuities. The human operator can attend closely to only one thing at a time; he is a single channel device in this sense. Tasks composed of a variety of activities require time to alternate attention from one activity to another. The driver sets up a steering response program that carries him for a few seconds, and can attend something else in that interim, e.g., tune in the radio or turn on the windshield wiper. This is the discontinuity of time-sharing. Displayed information is more often than not time-shared, in that a number of different information sources are inspected periodically.

In continuous operations such as tracking there is evidence that the human operator does not really function continuously, but that he formulates his response in segments averaging approximately 0.4 or 0.5 seconds. Part of the evidence for this is that the operator introduces these frequencies into tracking records even when they are absent in the curve he is following. Because of this discontinuity, some engineers have held that the human operator performs tracking operations more like a sampling servo with 2 or 3 per second sampling rate rather than like a continuous servo.

Yet another discontinuity or periodicity of the operator may exist in perception. It has been suggested that there is a process in visual perception that may be described as buffer storage of sensory data, which is scanned periodically in the perceptual process. The basic frequency of scanning might be about ten times per second, the frequency of the alpha rhythm, a primary rhythmic variation in the electrical activity of the brain.

These questions of periodicity of perception and response are potentially significant for the growth of an understanding of both perception and human movement. For purposes of this study we need only point out that it is quite possible that the perceptual and motor processes underlying human skills are rhythmic and periodic rather than continuous in nature.

VI. INFORMATION REQUIREMENTS AND RECEPTOR CHANNELS

Perhaps the central problem in manual control is that of presentation of information to the human operator. Certainly information display is the major avenue for improving most manual control systems. This chapter is concerned with the human operator's requirements for information in manual control and the sensory channels through which the information must reach him. Two subsequent chapters deal with displays.

A. The Information Requirements of the Human Operator

At any given hierarchical level in an automatic closed loop control process there are usually two and at most three types of information reaching the controller: input information, output information, and (sometimes) adaptive information, adjusting the way the controller responds to inputs and outputs. In the average automatic system, the "adaptive" input to the controller is no more than adjustments to the controller by a man, although in some systems adaptive loops are automated and keyed to, e.g., statistical functions of past input signals or sensed changes in the environment or system under control.

The human operator functions quite differently from the automatic controller. Because the operator normally does some planning, choosing the goal or the route to a designated goal, a function without counterpart in automatic systems, he requires planning information. Because he normally exercises control on the basis of the difference between what is desired or planned and what is predicted, he also has a need for predictive information. Planning and prediction are both oriented around the future, whereas information received by the operator, has reference to the past. The operator employs information about the past, then, to predict and to plan for the future. Only to the extent that there is prediction and planning is there control; the operator unable to look into the future is without control over it.

The requirements of the human operator in manual control are basically for predictive and planning information. Prediction and planning both involve the operator's internal model of the variable under control, however, so the information on which the model is built antedates predictive and planning information as such and will be discussed first. At the outset, information requirements for those man operated

systems in which there is an operator who does neither prediction nor planning will be briefly considered. The discussion of information requirements will therefore be presented under these four headings:

1. Command Information
2. Building an Internal Model; Adaptation
3. Predictive Information
4. Planning Information

1. Command Information

Man does sometimes serve as a control system element in roles in which he does little or no prediction or planning and is, in consequence, providing little or no control. This is the case in those tracking situations, for example, where the operator's problem is to follow a random input signal as precisely as possible. It is also true when the operator employs a "command" display which precomputes his desired response. The operator's task is in this case also tracking, but what is tracked is the output of a computing circuit. This computing circuit derives a desired operator response from input and feedback signals. Control in the case of the command display is taking place at a higher level of the control system hierarchy than that at which the tracking or command task is being performed.

Tracking can be regarded as one end of a continuum in manual control systems. As the operator is employed more for tracking and less for control, the requirement for planning and predictive information decreases, and human operations are oriented more about present time and less about the future. As the limiting case of no prediction and no planning is reached, the operator also exercises no control, but is serving only as a transmission link in the system. Interest becomes narrowed, in this case, to the transmission characteristics of the human tracking element. It is this limiting case that has been studied most extensively in research on human operator characteristics such as are reviewed in the Appendix. Since the primary concern of this report is manual control rather than signal transmission by the operator, our primary interest is in exactly those functions essential to control but irrelevant to such transmission. These functions are planning and prediction, and information that makes planning and prediction possible.

2. Building an Internal Model; Adaptation

Prediction and planning both involve the operator's internal model of the control process and variables in the environment that affect it. It is by means of the model that the operator is able to predict and to plan. Information to the manual control system operator is best understood in terms of how it contributes to the operator's internal model, and how the model is used in the prediction and planning processes. The operator's model for control has two aspects or parts:

a. Environment. Those external processes or variables that affect the controlled variable, and

b. System. The control system or process itself.

Both of these are modelled, in some systems the first being the more important for control, in some the second. In both cases there are three kinds of information of interest to the operator attempting to model the control system and its environment:

- Invariant features, as determined by unchanging structures behaving under fixed laws
- Adaptive features, referring to slowly changing structures or features
- Status information, referring to changing events, the state of the variable features of interest in the system or the environment.

To illustrate, the fixed structural features of an aircraft (e.g., shape, materials) and of its environment (gravitational field, characteristics of the atmosphere) determine invariant features that can be incorporated permanently into the pilot's model for controlling the aircraft. Adaptive features are not invariant but change slowly; for example, changes in aircraft response with atmospheric pressure and temperature change, or with weight changes due to fuel consumption. The physical laws governing these changes may be themselves invariant, but if the internal model employed for control does not incorporate the laws, then the operator may need to adjust his model "adaptively" in accordance with the slowly changing variable.

What is happening to the airplane, i.e., the moment to moment changes in heading, altitude, rate of climb, etc., comprise aircraft status information. Of the three different kinds of model information

important to the operator of a manual control system, only status information plays a major role in display. Slower changing and invariant features of a system do not lend themselves to displays of the usual sort. New kinds of displays that do indicate invariant and adaptive characteristics of a system are discussed in Chapter VIII.

The novice operating a manual control system for the first time has some understanding of the system, and some expectation, however crude or erroneous, as to how it will respond to his control actions. He thus begins operation with some kind of internal model of his system, which he employs to make predictions. The predictions he makes often prove in error, and force him to change his model. New predictions based on the revised model will usually be better but still in error, permitting additional, usually smaller, changes to be made in the model. As experience is gained, the model is gradually adjusted to further reduce errors in prediction. When the operator has learned the system and his performance no longer improves with practice, his model has stabilized, and errors in prediction have approached some asymptotic level.

The process of building an accurate internal model is the primary ingredient in training for manual control. By "accurate model" is meant one which serves as a basis for accurate predictions. It is the model that makes possible the predictions, and the error in prediction that makes possible the correction to the model. Thus the accuracy of the model depends upon the predictive process, and accuracy of the predictive process depends on the accuracy of the model, as we shall see.

The same process by means of which the model is developed and corrected is employed by the operator to make adaptive changes. When the predictions of a skilled operator begin to be in error, he can use this information to make changes in his model. He may or may not also infer the cause of the error in prediction (e.g., "my aircraft is getting out of trim"). The essential feature of adaptation is that the operator's model is changed, so that predictions based on the model and control activities based on the prediction reflect the adaptive changes.

3. Predictive Information

The operator's model is a structure built up in the past that he employs to predict the future. However, prediction requires not only that the operator have some workable model of the control process; it requires also that he be aware of the current situation. He needs to

know not only the invariant and present adaptive structural features of his model; he needs to know also what is happening, what are the present and changing events which he apprehends (models) in terms of the status or state of his system and its environment.

The usual form of display for manual control is the status display, which shows the operator the state of his controlled variable and other variables affecting it. This category includes such instruments as compasses, altimeters, pressure gages, speedometers, temperature indicators, air speed indicators, flowmeters, submarine depth gages and pitch angle indicators, rudder angle indicators, and in fact almost any display that is named. Status information is the kind of information that the human operator receives directly by observing changes in his environment when no display intervenes. In this sense the status information is the most fundamental type of information for manual control. Yet status information in itself has absolutely no value to the operator for, as has been said before, it refers to the past, to what has already happened. Since finite time is required for any information to reach the operator, it follows that everything he receives refers to what has gone before. Since the past cannot be modified, the operator has no control over it. The operator is not interested in status information as such, then, but only because he can use it to predict.

Prediction is accomplished by the operator by combining status information into his model. Because his conscious processes are not limited to present time, the operator can extrapolate status information into the future through his model. The basic prediction takes the form, "If this is the way the system works (the operator's model) and this is what is now happening (status information) then this is what is going to happen unless I do something to change it". More sophisticated predictions which include the consequences of alternative control actions available to the operator are employed in planning.

a. Derivative Information in Prediction. Derivative information has particular usefulness for prediction in high order manual control systems, especially vehicles. When a controlled variable moves through space or changes with time in a continuous fashion, and when its space or time derivatives can be measured and displayed to an operator, this can be of great predictive value to him. Derivative information is status information, of course, and provides a means for more accurate model predictions. In some higher order systems, however, the operator forms another model, a "derivative model" which he employs for prediction. This model is likely to be built around the perception of the derivative display. Consider air speed in flying: the

operator thinks of air speed, not just in terms of the motion of his craft through the air, but also (and especially, for purposes of control) in terms of the position of the pointer on the airspeed indicator. The derivative model allows a rate or an acceleration to be perceived and represented conceptually as a position.

Since many physical laws are expressed in the form of differential equations, the derivative model may relate to the applicable law or to the structure or functioning of the system controlled in a way that is simple and logical, i.e., the derivative model may behave in a way that is easier to understand than does the controlled variable. -- And the derivative model usually bears a clear and straightforward relation to the controlled variable, so that it can be employed effectively to control the latter.

An operator's derivative model, like his model of the controlled variable itself, has invariant and adaptive features which the operator learns over time. Damping or unstabilizing feedback effects, for example, affect the response characteristics of the derivative model, and must be learned. Prediction then is not a matter of simply extrapolating derivative information that might appear on an indicator, but extrapolating it through the operator's internal model. The model incorporates the response characteristics of the controlled variable, or whatever function related to the controlled variable, the derivative of which is displayed.

b. Input Prediction. Prediction in some manual control systems is concerned just with the controlled variable and variables which affect it. When the input is constant then the problem of control centers on the output and factors which may cause it to change. The aircraft travelling at constant altitude, the room at constant temperature, the ship sailing at constant heading, the satellite maintaining a constant attitude, and the chemicals being kept at constant pH are all examples of systems in which the input is usually constant, and prediction is concerned with the output. There are also systems, however, in which the input varies, and much or most of the problem of control is concerned with the input. Consider radar tracking systems for aircraft, the control problem of the anti-tank gunner, the altitude control problem of the pilot who must fly near the ground over rough terrain, or even the steering problem of the automobile driver on a winding mountain road. In all such cases the input, the desired value of the controlled variable, is tied to some variable in the environment. In these cases the prediction of input, of what is going to be desired,is the primary concern in control, and the predictive process focusses on the input. The operator's model of the controlled variable then takes a back seat to his

model of the aspect of the environment to which the input is tied. Predictive information about an environmental variable to be followed is usually quite different in character from that about the control system or process. Any kind of information that enables the operator to anticipate his input is enormously helpful. The prediction involved in input anticipation has been studied quite a bit, especially by Poulton[1] and by North[2] in England, and by Krendel and McRuer[3] and Sheridan[4] in the United States. Information predictive of the input is sometimes referred to as "precognitive information".

"Precognitive information" may be directly available, i.e., without the necessity for a prediction by the operator, as when the director sees the road ahead or the pilot views a radar display of the terrain. In other cases, the operator's model of the environment and status information about it are used to extrapolate, much as with output prediction. In a third and particularly interesting case, the operator shows the ability to form a different kind of model -- a stochastic or probabalistic model, which he uses to make a statistical prediction of input signals having a random character, but statistically definable amplitude and frequency characteristics. An early study by Ellson and Wheeler established that the operator does respond to statistical properties of his input signal.[5] Poulton explored this function

---

[1] Poulton, E. C. Perceptual Anticipation in Tracking. Cambridge, England: Medical Research Council, Applied Psychology Unit Report 118/50, 1950; and
Perceptual anticipation in tracking with two-pointer and one-pointer displays. British Journal of Psychology, July 1952, 43, pp. 222-229.

[2] North, J. D. The Rational Behavior of Mechanically Extended Man. Paper presented at Shrivenham, England: Conf. at the Military College of Science, 1954.

[3] Krendel, E. S., & McRuer, D. T. A servo-mechanisms approach to skill development. Journal of the Franklin Institute, 1960, 269, pp. 24-42.

[4] Sheridan, T. B. On precognition and planning ahead in manual control. Washington, D.C.: Fourth National Symposium on Human Factors in Electronics, May 1963.

[5] Ellson, D. G., & Wheeler, L. The Range Effect. Ohio: Wright-Patterson Air Force Base, Air Materiel Command Technical Report 5813, May 1949. ATI 53 593.

more fully.[1] There is no question but that an operator tracking a "random" input signal does, in fact, learn and use statistical properties of an input signal, and use this predictive information in his response. The information that a signal is more likely to be in one direction than another or of one amplitude than another is of some help to the operator endeavoring to pre-plan his response.

It appears that predictive information is so important in control tasks that the operator uses almost every conceivable means for obtaining it. Predictive information is normally generated inside the operator as a consequence of projecting status information through some kind of model formed by the operator. Chapter VIII, however, describes a technique for displaying such predictive information directly.

4. Planning Information

When an operator's task is completely specified for him, as when he follows the output of a command display, he has no need for planning information. The display specifies the operator's response, which is pre-planned to generate the display signal. As has been indicated, this situation uses the operator as a transmission rather than as a control element. Ordinarily the human operator has some range of choice, however, even in what are ostensibly command systems. When there is a choice, when the operator selects alternatives, then he is truly exercising control, and he has a requirement for planning or goal selection information.

Planning requires two kinds of information -- predictive information and evaluation information. The predictive information needed is of the sophisticated type associated with the operator's control alternatives. The prediction of the consequences of alternative actions available to the operator is usually the most difficult part of planning.

Given the predicted effects or consequences of some of the alternative control actions available to him, the planner still must make his choice among the alternatives. His choice may be influenced by many different factors. There is frequently a defined goal from a higher hierarchical level that serves as a primary constraint on the operator's choice. The

---

[1] Poulton, E. C. Learning the statistical properties of the input in pursuit tracking. Journal of Experimental Psychology, July 1957, 54, pp. 28-32; and
On prediction in skilled movements. Psychological Bulletin, 1957, 54, pp. 467-478.

lower level loop is at the service of the higher, and the choice made at the lower level must fill the requirements of the upper. For example, the ocean liner captain may have a range of headings to select from in planning his course, but they must all lead to his destination.

The higher level goal may be embodied in a definite input signal to the operator. The command of the captain of the ocean liner to his helmsman to steer 85 degrees, is an example. To the policeman following a suspect in traffic, the other car's course defines an input signal. The operator of a power generating plant may use a display of power being consumed as an input signal for his control over steam pressure; this display, then, represents an input requirement of the system.

When there is no higher level in a particular control system hierarchy, i.e., when the operator does the planning in the outer loop, he has no input signal, but formulates the system goal as well as the route to reach it. The man going where he chooses in his automobile illustrates the point. There does not always have to be an input signal.

Inputs defined by a higher level loop in a control system hierarchy are a first source of planning information to a human operator, then, but there may or may not be a higher level loop. When there is, then the operator's plans are limited to those which will lead to the specified goal; when there is none the operator has a wider range of freedom to plan, as he may select the goal as well as the route to it. In either case, he still has alternatives to select from in planning. These alternatives may be further narrowed by applying criterion information.

Criterion Information and Planning. While goals from a higher level in a control system form input information, other goals than these may be applicable as criteria for choosing among control alternatives. The requirement is that the alternatives available to the operator are predicted to be differently effective in achieving or advancing toward the goal represented by the criterion. The planner in receipt of this information can then use this criterion information in addition to input information in making his choice.

Criteria vary according to whether they are integral to the operation of a system or external to it. Performance criteria are integral, and are concerned with such things as how quickly and precisely a higher level goal (input) is achieved. External criteria refer to different goals, however, that may be differentially affected according to the

alternatives chosen. One choice may be safer or more economical than another, for example. Among the criteria external to the operation of the system are also the personal inclinations of the operator, who may find one alternative, for instance, easier or more interesting than others.

The use of criterion information imposes an additional requirement on the internal modelling process of the operator. Not only must his internal model represent the controlled variable and factors related to it, but it must also incorporate the relations of the controlled variable to criterion variables. This may multiply the modelling information the operator requires for control.

The expanded model which incorporates criterion variables is also employed predictively, for criterion information, too, is directed toward the future. The operator's choice is influenced by how he predicts that various control alternatives will affect those criteria he is employing.

B. Information Channels of the Operator

Perception forms the means by which we become informed about the world around us. The senses are special channels that provide the information we need to build our model of the world, and to formulate possible modifications in it. Each sense modality should be regarded as an avenue of information to the operator, information which takes the form of characteristics of the operator's model. Each sense contributes to the model through its own special building material, i.e., the sensory qualities associated with that sense. The model-making activity is, nonetheless, a unitary process, whatever materials are employed in the construction.

The senses primarily involved in manual control are, in ascending hierarchical sequence, i.e., from inner loop to outer:

- tactual-kinesthetic senses
- balance
- hearing
- vision

Some of the characteristics of each of these senses or groups of senses will be discussed briefly.

1. Tactual-Kinesthetic Senses

Our internal model of the world is built up from movement through, and manipulation of objects in, the environment. The basic senses, both ontogenetically and phylogenetically, are those which underlie movement. They are the inner loop of the control process, and are the basis for the outer loop senses that make possible longer range planning. The latter are, of course, a later development. The child responds to touch before he responds to vision.

The tactual-kinesthetic senses respond to pressure and temperature of objects on the skin, to joint and muscle tension, and certain other stimuli. Their combined effect is to enable man to move and perform manipulations skillfully. Through them an operator is able to position a limb or a control operated by a limb accurately, or to apply a desired amount of force to a control. Disruption or interruption of the internal feedback loops carrying kinesthetic information results in gross disturbances in the ability to perform "simple" movements like reaching and grasping a glass of water, holding it upright without watching it, and drinking from it without controlling the movement visually.

2. Acceleration Senses; Balance

A group of senses cooperate in supplying information about accelerations that are applied to the whole body. Such accelerations result in stresses on the body to which the tactual-kinesthetic senses respond. The "postural" reflexes necessary for body orientation and coordinated movement are based on these, and on the sense of balance.

The sense of balance, which is based on a highly developed sense organ in the inner ear, responds to acceleration of the body. Although usually discussed in regard to the acceleration of gravity, the sense actually responds to the resultant of overall accelerative forces. Expressed differently, field accelerations, whether gravitational or inertial in origin, result in a force acting on each particle of the body. Unless the body is in free-fall, its position is restrained by something. The pushing against the external restraint is perceived through the tactual-kinesthetic senses. The sense of balance receives information from the internal "pushing" of an especial sense organ that indicates the direction of the overall bodily acceleration with respect to which the body is restrained. Were there no restraint, of course, the body would actually be in free-fall, and the sense of balance could not perform its normal function.

3. Hearing

Tactual-kinesthetic senses depend on bodily contact for sensory excitation; acceleration senses depend on it to the extent that the body must be restrained. Neither involve distance reception, as do hearing and vision. Hearing is responsive to the mechanical vibration we call sound, which may be transmitted by both fluid and solid media. The air carries and diffuses sound waves to the highly refined apparatus of the ears which converts these waves to nerve impulses. The diffusing of sound waves results in some of the most important characteristics of hearing as opposed to vision. Major among these is that sounds do not convey accurate information about the shapes of objects; hearing is not a spatial sense in the way that vision is. Sound signals carry information as to the location of the sound source, but not as to its fine spatial structure. Sounds are not finely patterned in space in the way light is, and to convey information the signal is instead finely structured in time. This requires that it be perceived sequentially over time. The ears are extraordinarily sensitive to temporal structure, and so hearing is the primary communication channel for information passing between people.

Conscious processes in general are structured also primarily in time, so that the limitation of sound in transmitting information is much less than would otherwise be true. Vision is truly a spatial sense, but written information, in the form of the printed page, is read sequentially a few words at a time, and in the process translated into something much more akin to speech communication in order that it be understood.

Hearing provides the feedback loop to monitor what is very likely the most exacting motor task there is: the playing of a musical instrument. There is probably no control system in the world which requires of a human operator anything approaching the exceedingly precise positioning reactions made each second by the violinist's fingers, while the sequence of lever operations of the skilled pianist, the sequence and timing and the hardness or softness of each stroke make the task one of fantastic difficulty compared with most human performances. It is difficult to say whether control systems will ever be able to utilize the full range of skill that musical performance demonstrates is within the capability of man to achieve.

At present, hearing is useful in control systems primarily for speech communication, and for display of signals requiring immediate attention. Since the ears cannot exclude sounds, hearing is the natural

choice for emergency signals. There have been a few interesting attempts to utilize hearing for display to an operator of tracking type information, but these have never become widespread.[1]

4. Vision

In most control systems, vision is the primary channel for carrying information to the operator. Like the ears, the eyes are distance receptors, their adequate stimulus being patterns of electromagnetic energy of appropriate wave length composition. This energy travels in straight lines, so that reflected light, unlike reflected sound, preserves a formal spatial correspondence to surfaces it strikes.

The retina of the eye is a surface onto which images formed (mostly) by reflected light are focussed. Unlike hearing, the spatial structure of these images is preserved in great detail, the eye thus being able to receive a substantial amount of information all at once and non-sequentially. However, the acuity of vision is great only in the center of the field. Patterns of any complexity are not perceived all at once but in sequence. We do not fixate a picture in the middle and abosrb it all, but let our attention -- and eyes -- move from detail to detail. However, the internal model preserves a correspondence to the spatial structure of the object viewed that is independent of the sequence of fixations. The small segments of spatially structured sequential information that we receive through successive eye fixations form this larger coherent model, and the model preserves a spatial correspondence to the environment that goes far beyond the spatial information received in any one fixation.

The selective aspect of vision is another important point of difference from the sense of hearing. The eye is mobile, and in being directed at one part of the visual field excludes the remainder. This contrasts with hearing, of course, as the ears have little capacity for selecting among sound signals. For this reason, the eyes are the natural channel for displaying information that does not require continuous or immediate attention.

---

[1] Forbes, T. W. Auditory signals for instrument flying. Journal of Aeronautical Sciences, 1946, 13, pp. 255-258.

5. Vision and the Information Rate of the Human Operator

Perception is a selective, filtering process, involving a great reduction in information. This can best be illustrated with respect to vision, where the reduction is most striking. The reduction begins with the selection of only a small portion of the available visual stimuli for attention. This results in a selected image pattern falling on the retina, the retina being a mosaic of more than a hundred million individual receptor cells. Since each receptor can, under appropriate conditions, produce a perceptable response, the information capacity of the retinal surface itself might be computed by multiplying the number of receptors times the rate at which any one receptor can respond temporally to intermittent stimulation. This is on the order of ten times per second. Thus the retinal surface might be said to have an information rate on the order of a billion bits per second. Receptors converge onto fewer nerve fibers, so that the optic nerve contains only on the order of a hundred thousand neurons, and might be saia to have a potential information rate of some one million bits per second. -- Yet the rate at which information received through the eyes is utilized in perception is probably no more than about a hundred bits per second. Such a figure might obtain for reading, for example.

In a typical motor task such as typing, the information rate through the operator is on the order of twenty to thirty bits per second. In continuous control in one dimension, the operator can track random signal frequencies going to up to about three cycles per second, with an accuracy corresponding to an information rate of six to eight bits per second.

The ability of the human operator to transmit information is thus extremely limited. In engineering parlance, the operator cannot be used as a series element in a system requiring a high bandwidth or high information rates. The virtues of man as a control system element rest on considerations other than his quantitative information transmission ability; it is instead one of his fundamental limitations.

VII. DISPLAYS

The function of displays in manual control is to provide the operator with information he requires to exercise control. This is the information he needs to predict the consequences of control alternatives available to him, and to evaluate them and plan accordingly. The operator perceives, predicts, and evaluates on the basis of what we have called an internal model of the controlled variable and factors relating to it, including criterion factors. Displays other than the command display are designed to enable the operator to build an effective internal model, which in turn makes possible effective prediction and planning.

Channels for presenting information to an operator define one important set of limits in the design of displays. Another important set of limits is determined by what can be sensed, as discussed below. Within these limits there are numberless ways in which displays can be designed, only a few major aspects of which can be covered in this brief survey, which will center about presentation of visual information in analog form. The principal topics to be taken up are:

- A. Sensors for Displays
- B. The Analog Display
- C. Display Coordination and Integration

Chapter VIII which follows deals with some important special display techniques.

A. Sensors for Displays

When an operator exercises control via displays rather than direct sensory contact with his environment, the information must be obtained to appear on the display. A sensing device of some sort is employed to obtain the information. What can be sensed forms a fundamental limiting feature of displays. This limiting feature is not always given the emphasis it deserves. To repeat an example I gave of this point many years ago:

> A measuring instrument which indicated directly and accurately its position relative to the earth's surface would render obsolete the entire science of navigation as we know it ....[1]

This one example illustrates clearly the limits imposed on navigation systems by lack of a sensing instrument. The whole system of celestial and radio fixes, the use of compass information, inertial and other "dead reckoning" systems all are substitutes for what we would like to sense directly and display, but do not yet have the sensing means for.

The initial step in considering the design of displays for a particular manual control system is to analyze the information the operator would really like to have, and to consider the sensing instruments available to obtain it for him. Too often this analysis is not carried out, and it is assumed that the operator requires information that, e.g., has been displayed on similar systems in the past. It is likely that no analysis was carried out for the previous systems either, and with the developments that have taken place in the field of instrumentation, an appropriate sensing device may have been developed where none formerly existed. The designer for automatic control systems knows that sensors are more often than not his most difficult problem. In manual control there is not always the same awareness of the problem.

B. The Analog Display

1. Analog vs. Symbolic Displays

Displays convey information to a human operator in two different ways. Symbolic displays code information in an arbitrary and conventional manner, employing, e.g., words, numbers, or colors to represent aspects of the environment. The symbolic display can have no meaning to the person unfamiliar with the code or convention. One who is familiar with the code, however, is able to interpret the information which, in a control system context, means that the operator can use the symbols displayed to refine his dynamic internal model of some aspects of the environment related to the variable under control. The process of interpretation is essential to the symbolic display.

---

[1] Kelley, C. R. Submarine Control by a Single Operator. Port Washington, N. Y.: U.S. Navy Special Devices Center Technical Report 954-00-18, October 1953, p. 23.

Displays are discussed later in terms of the concept of "display space", defined as the family of discriminably different display configurations. The symbolic display has an enormously expanded "display space" compared with the analog. The pointer display with a three inch scale might have 30 discriminably different display configurations when used on a manual control panel, while the symbolic display occupying the same space might easily have 30 billion!

The analog display is a physical model of something. It may be greatly condensed, transformed, etc., but it retains some formal correspondence with the thing displayed. Instead of representing something by arbitrary and conventional rules of correspondence, the analog display represents by means of a physical transformation. Commonly the correspondence is spatial, and the position and movement of analog display elements correspond directly to the position and movement of something which the display represents. If the spatial correspondence is such that the display resembles that which it represents, the display is said to be "pictorial". Pictorial displays require a minimum of interpretation, as the nature of the analogy is immediately obvious.

Analog displays are by no means limited to spatial representations of spatial variables. Non-spatial variables such as temperature, voltage, pH, and hundreds of others are frequently represented on a display by a spatial analogy. There are also non-spatial analog displays, such as the army's doppler radar, by means of which an observer can hear signals indicating the movement of vehicles and personnel via an electromagnetic signal converted to a sound pattern. Sonar forms another example of an auditory analog display. (A symbolic auditory display might instead use words or numbers.) Most analog displays are visual, however, and our principal concern here is with their visual characteristics.

The distinction between analog and symbolic displays is itself not always clear-cut, and in addition many displays incorporate both analog and digital information. Baker,[1] for example, refers to moving-pointer displays as symbolic. When the term "symbolic" is used this broadly, it includes both analog and symbolic displays as defined here.

---

[1] Baker, C. A., & Grether, W. F. Visual Presentation of Information. Ohio: Wright-Patterson Air Force Base, Air Development Center, Technical Report 54-160, August 1954. AD 43 064.

All displays operate on analog principles, in the sense that there is some correspondence between the changing element of the display and that which the display represents. However, the symbolic display presents the correspondence via an arbitrary code rather than a physical analogy. Figure VII-1 is a series of hypothetical pitch angle displays showing the angle of dive of a submarine. All of these indicators operate by means of elements that move in correspondence to pitch motions of the submarine. The displays are increasingly symbolic in emphasis, although all have both analog and symbolic elements. To the extent that a submarine planesman might obtain information for control directly from the position and motion of the display elements, he would employ the analog elements of the display; to the extent that he used the numbers he would employ the symbolic.

Many displays for manual control take the form of analog displays to which symbols have been added. It appears that such displays give an operator order of magnitude and/or rate of change information by their analog structure, and quantitative information symbolically.

Alphanumeric symbols are by nature discrete, and it is generally but not invariably true that symbolic displays are discrete and analog are continuous. Analog displays may also be discrete, however, especially when they represent discrete variables such as switch positions.

Because of the very limited quantitity of information available on the analog display as opposed to a symbolic display of the same size (a difference illustrated above by the figures 30 vs. 30 billion discriminably different display configurations) it might be wondered why analog displays are employed at all. In fact, the great majority of displays for manual control are of the analog variety. Since the information transmission rate of the human operator is low, the information transmission rate of his display is likely to be secondary to other considerations. Since the human operator exercises control by reference to an internal model, information displayed to him in a form compatible or assimilable with his internal model is likely to be of most value to him. This accounts for the value of the analog as opposed to the purely symbolic display.

2. Display Motion

Displays present signals which change with time. An unvarying signal would carry no information and require no display. The way in which a display indicates change is a fundamental property, often more

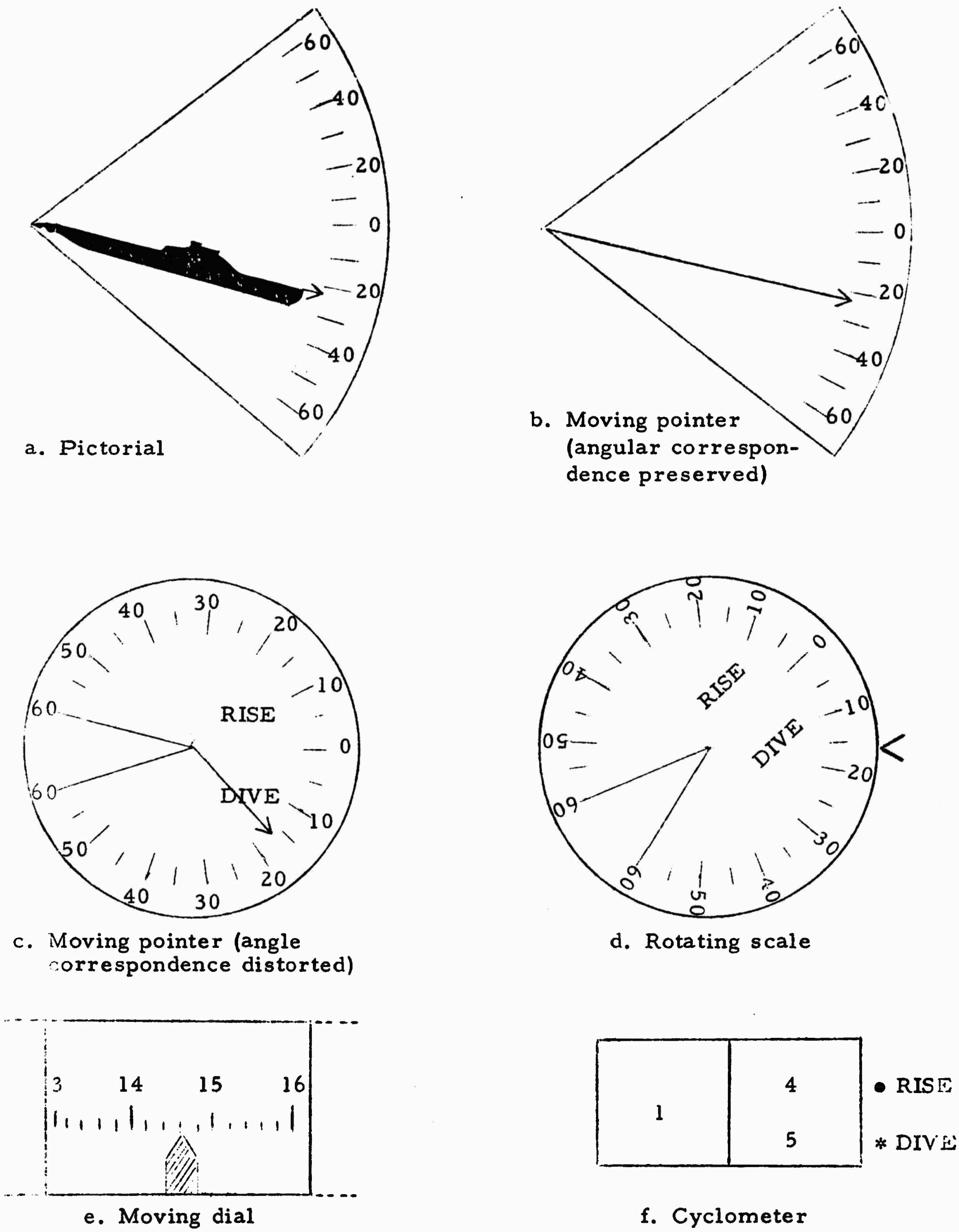


Figure VII-1. Analog vs. symbolic displays of submarine pitch angle. The last two are virtually purely symbolic.

important than the way it indicates quantities or spatial relations. Displays may indicate change by variations in brightness, color, etc., but the primary indication is by motion of a display element. The moving display element is to the analog display what alphanumeric symbols are to the symbolic display.

All motions are relative to something. The moving display element moves with respect to an indicator background or scale, or the display panel around the displays or even the operator himself. Unless the rate of motion is very slow the operator perceives the moving element directly as moving, while the background serves as the frame of reference for this motion. In the case of very slow motion, motion in itself is not perceived, but rather differences are detected in successive observations of position, and motion is inferred. The second hand of a watch is observed to move; the minute hand is inferred to move.

Under exceptional circumstances the moving element of a display can become the frame of reference for motion, such that the background and surround appear to move and the display element is perceived as being at rest. This requires a large and compelling display. A horizon display subtending sixty-seven degrees of visual angle was not enough in one study.[1] The size of the screen in Cinerama (as opposed to smaller "wide-screen" motion pictures) and its effectiveness in creating the illusion of observer motion illustrates the display size here discussed. When motion cues other than vision are present which indicate to the observer that he is moving with the display element, it can help to make the display element appear as the frame of reference with large articulated displays. With displays of normal size, however, it can be assumed that the display background will form the frame of reference to which motions of the display element are referred.

3. Motion Compatibility

The direction in which display elements move in relation to the movement of objects represented by the display and to associated controls is the problem of motion compatibility. The simplest case of motion compatibility is the command display, for here the display refers directly to the space in which the control moves, i.e., the display tells the operator what to do with his control. For a command display showing ordered and actual control position separately, e.g., via two pointers

---

[1] Kelley, C. R., De Groot, S., & Bowen, H. M. Relative Motion: III. Some Relative Motion Problems in Aviation. Port Washington, N. Y.: U.S. Naval Training Devices Center Technical Report 316-2, January 1961.

(i.e., "pursuit" displays) there is no problem. The control and the display of actual control position should both move in the same direction, insofar as this is possible. The command display combining desired and actual control position on a single display of error ("compensation" display) in control position is less simple, for depending on which one of the two display elements is considered the frame of reference, the display may move in the same or in opposite directions relative to the control. However, the operator using a single element display will spontaneously see his display, not as a command to move his control in the direction the display element moves, but as something he has control over which he should bring back to center by moving his control in the opposite direction. This serves to define the appropriate display configuration.

With displays other than command displays compatibility problems increase, because these displays refer, not to the position of a control, but to other aspects of the system which may also move in space. Thus it is not a question of the relation of display to control, but of display and control to a controlled element, as well as to each other. Evidently, if a control moves, e.g., up or to the right, it should produce a corresponding direction of motion in the controlled element, which should be indicated by a corresponding motion of a display element. Compromises are necessary, of course, due to the fact that the planes of motion of the three elements may be different. There is a fundamental ambiguity, for example, when a fore-aft motion results in an up-down motion, or vice versa, to which no simple compatibility rule can be applied.

Even more serious problems occur in the case of moving vehicles because of the frame of reference problem. If the vehicle operator considers his vehicle the frame of reference and the environment about it as moving, displays ought to move in the opposite direction than if he were to consider the environment as the frame of reference and the vehicle moving. The former is referred to in the literature as the "inside-out" frame of reference, the latter the "outside-in". Here we will call them vehicle coordinates (inside-out reference) vs. external coordinates (outside-in reference). In some cases one frame of reference appears appropriate, in others the other, and many ambiguous cases exist.

Reference for attitude information concerns the orientation of the frame of reference, and reference for position information concerns the location of the origin of the coordinate system for the frame of reference, a separate problem. Either attitude or position can be referenced externally or to the vehicle. Thus when attitude and position information

are both displayed to a vehicle operator, there are four display reference possibilities:

- externally oriented coordinates with fixed external origin;
- externally oriented coordinates with the vehicle the origin;
- vehicle oriented coordinates with externally fixed origin;
- vehicle oriented coordinates with the vehicle the origin.

To apply this to a specific case, consider the problem of a dynamic map type display for, e.g., aircraft maneuvers in a crowded airspace, or other vehicle applications. The display will show both position and heading by means of a vehicle symbol on the map, but may do this in any of four ways, corresponding to the above display reference possibilities:

- north-oriented stationary map, moving and turning symbol;
- north-oriented moving map, turning symbol;
- turning map, moving but unturning symbol;
- moving and turning map, stationary symbol.

Every possibility has display problems. Upside down maps are hard to read, while the vehicle symbol moving south on a north-oriented map is 180° out of orientation with the real vehicle, so that a left movement of the control brings about a movement of the vehicle symbol to the right. The display must jump to show a new section of map when a moving vehicle symbol reaches the edge of the map; the jump frequency will depend on scale, and can be a problem. -- And both the moving and the turning map involve the fundamental problem of display motion compatibility, which is that the operator tends naturally to perceive a moving display as a controlled element moving with reference to the display surround, i.e., the panel. He does not normally see the display element as the reference, and the panel and other parts of the vehicle as moving with respect to the display, regardless of whether or not the vehicle is "really" the moving element, and the display element "really" stationary.

All displays of fixed external references as moving elements appear to a vehicle operator to move the wrong way, and create serious problems of motion compatibility. The aircraft artificial horizon display is the

most studied example of this problem.[1,2] Artificial horizon displays, like all vehicle displays which show fixed external references as moving display elements, appear to the untrained operator to move in the wrong way. Although an operator can be trained to respond correctly to such displays, he is not trained out of the perceptual factors which determine for him what appears as frame of reference and what appears as moving. In periods of forgetfulness or disorganization due to stress there is always the danger that the operator may revert to his natural mode of responding, and move his control in the wrong direction. In one well-controlled experiment in an aircraft fire control simulator, moving horizon displays resulted in ten to twenty times as many control reversal errors as moving aircraft-fixed horizon displays.[2]

The use of an external reference of some sort as the moving element for a vehicle display often finds strong supporters among engineers. After all, the operator does maneuver well by such reference through windows and windscreens. When the visual display cannot be made as large and compelling as the pilot's view through the canopy or as, say, Cinerama, then motion compatibility problems arise, and serious consideration should be given to alternative means for presenting the required information. The alternatives available invariably have disadvantages also. In a trade-off analysis the fact that motion incompatibility will at times bring about wrong movements of a control in even a highly trained operator should never be neglected if such wrong movements can have serious consequences. In vehicle cases they usually can.

The usual alternative to displaying external reference information as the (apparently) moving elements of displays is to show a reference fixed with respect to the panel with a moving element representing the vehicle. This form of display is intuitively unappealing to many because it does not show "true" reference information at all, and because it displays the vehicle symbol as if the vehicle were viewed by an outside observer anchored to the fixed reference, while the operator is actually inside the vehicle. These objections lose some of their force on analysis. That the operator is located in the vehicle is in itself less important than the kind of internal model he has built of the vehicle situation. If this

---

[1] Fitts, P. M., & Jones, R. E. Reduction of pilot error by design of aircraft controls. Air Technical Intelligence Technical Data Digest, 1947, 12, pp. 7 - 20.

[2] Bauerschmidt, D. K., & Roscoe, S. N. A comparative evaluation of a pursuit moving-airplane steering display. IRE Trans. of the Professional Group on Human Factors in Electronics, September 1960, HFE-1, pp. 62-66.

model is of a fixed reference frame and a moving vehicle (his), then the external reference outside-in moving vehicle display is directly compatible with this internal model, and the moving reference display is not. For example, the pilot of the banking aircraft is likely to have an internal model of a banking aircraft over a fixed ground reference and not the slanting ground that he actually sees outside his canopy. A display of a banking aircraft is therefore compatible with his model, and from a psychological viewpoint a better choice than a moving horizon display. In other cases the operator may adopt a vehicle reference framework for his internal model at least some of the time, in which case an externally referenced display is more defensible. The moving map display may be a case in point here.

Compromises between vehicle reference and external referencing of displays are possible and, in cases where operators have been trained on "wrong" displays, maybe desirable. Several years ago Fogel[1] proposed a horizon display which provided an attractive solution to the problem of the moving horizon vs. moving aircraft display. The "Kinelog" is a display with a moving aircraft symbol that moves quickly and in the correct direction in response to the control, and a moving horizon that reorients slowly, so that in a sustained bank the aircraft symbol returns gradually to level as the horzion symbol assumes the angle of the true horizon. By separating out the high frequency aircraft response important for motion compatibility from the low frequency horizon reorientation, the Fogel display succeeds in presenting a display of a moving reference element that still moves compatibly when changes in bank angle are made. The display is, according to Fogel, also in good agreement with the way the pilot perceives bank kinesthetically. This frequency separation principle can be applied to other cases where there is some reason to retain the externally referenced display element.

4. The Display Space

The positional and metric properties of analog displays are likely to be less important in manual control than are movement characteristics, although they have received more attention. Many good studies of the

---

[1] Fogel, L. J. A new concept: The Kinelog display system. *Human Factors*, April 1959, 1(2), pp. 30-37; and
Fogel, L. J. *Biotechnology: Concepts and Applications, Chapter 9: Manual Tracking Decision*. Englewood Cliffs: Prentice-Hall, 1963.

accuracy of scale reading as a function of indicator design variables have been made, for example. These are reviewed in publications such as Baker and Grether[1] and McCormick[2] and will not be discussed here. This discussion will be restricted to problems of display reference and display scaling.

Analog displays for manual control commonly provide information about location or orientation, or changes in location or orientation, (i.e., translation and rotation) of a controlled element. In command displays they appear instead in terms of desired vs. actual positions of the operator's control. In any case, the space of the display represents some different region of space, and the changing display element conveys the required location, orientation, or motion information. The region of space represented on the display may be real or predicted, or may have only an analogical existence like the hyperspaces of physics. Displays of non-spatial quantities represent analog rather than real regions of space. Real spaces present the more serious display problem. However, it should be mentioned in connection with the display of non-spatial variables or dimensions that the culture has established "pupulation stereotypes" of display such that up, to the right, and clockwise directions are associated with increases. When the reverse direction of increase is employed in a display, it is read more slowly and mistakes are more likely.

Position, orientation, and motion must be defined with respect to something, the "something" being the frame of reference. The natural frame of reference for display motions has already been discussed in some detail. It is defined by the background and surround of the moving element of the display. The frame of reference for position and/or orientation is closely related, but not the same, for it is determined by elements of the display itself. The scale defines the frame of reference for pointer type displays, for example. The two dimensions of the cathode ray tube face define the frame of reference for two dimensional displays. A third and analog dimension maybe added to the CRT display by perspective and other coding methods to create the illusion of depth or distance. Of course, artists have used the perspective technique for hundreds of years to create a three dimensional frame of reference for positions and orientations of objects. Additional analog "dimensions" of a non-spatial character can be created by various coding schemes, employing such variables as shape, brightness, or color.

---

[1] Baker, C. A., & Grether, W. F. Visual Presentation of Information. Ohio: Wright-Patterson Air Force Base, Air Development Center Technical Report 54-160, August 1954.

[2] McCormick, E. J. Human Engineering, Chapters 11-14. New York: McGraw-Hill, 1957.

5. Display Scaling Techniques

It is useful to define display capacity for quantitative information in terms of "range number", the total of discriminably different display configurations. This definition has the advantage that it can be applied alike to real and analog display dimensions, and provides a basis for metric comparison of qualitatively different display "dimensions", including the "dimensions" of non-visual displays. Range number should be computed, not in terms of laboratory threshold data, but in terms of practicably useful discriminations for a working system. The range number of a three inch scale length, pointer-type display at normal panel distance is only about 20 or 30 at best, and with poor indicator design may be only half that many. Threshold detection data would imply a number closer to 100, but under practical normal conditions of slant, parallax and illumination, the figure 20-30 is not overly conservative. As has been mentioned, a symbolic display occupying the same panel space would have a range number in or beyond the billions; that is, it could show billions of alternatives alphanumerically. Thus the analog display has an extremely limited range number compared with the symbolic display.

For an equal-interval display scale with discrimination the same throughout the range, display range number is range divided by reading accuracy, both expressed in standard display units. A rudder angle indicator with a range of $\pm 27^{\circ}$ which may normally be read to within 3° has a range number of 54/3 or 18. Under conditions where a 1/8 inch separation in pointer positions is required to assure differential readings, a display with a range number of 18 must have 2-1/4 inches of scale. Table 1 illustrates range numbers for some representative displays.

Given the maximum interval size needed for discrimination, this can be multiplied by the range number to determine the scale length needed. The interval will, for practical purposes, vary from about 1/20" to 1/4" for 20 to 36 inch panel distances, with an average range of illuminations and viewing angles, and reasonably good display design. Note that employing 1/10" as the interval size needed, the range numbers of Table 1 would result in displays having scales ranging from .8 inches up to 1800 feet in length. Many displays obviously cannot be shown in analog form on a single continuous scale.

There are five principal techniques for dealing with the limited scale of the analog display: the vernier (multiple) display, multiple element display, multiple dimension display, partially symbolic presentation, and nonlinear scaling. All but the last of these are based on the same

Table 1

Approximate Range Numbers of Some Representative Displays

Display	Total (assumed) Range	Assumed Accuracy	Range Number
Fuel Gage (automobile)	0 - 15 gal	± 1-1/2 gal	10
Speedometer (automobile)	0 - 120 mph.	± 5 mph.	24
A. M. Radio Dial	550 - 1600 kc.	± 10 kc.	105
Ruler	0 - 12 "	± 1/16"	192
Thermometer	- 20 - + 120°	± 1/2°	280
Protractor	0 - 180°	± .5°	360
Pitch Angle (submarine)	± 45°	± 1/8°	720
Ship's Compass	0 - 360°	± 1/4°	1,440
Depth Gage (submarine)	Keel to 800'	± 1/2'	~ 1,600
Altimeter (aircraft)	0 - 50,000'	± 10'	5,000
Clock	12 hrs. (43,200 sec.)	± 1 sec.	43,200
Electric Meter	0 - 100,000 kw.hrs.	± 1 kwh	100,000
Spacecraft Sextant Readout (star angle)	± 90°	± 3" arc	216,000

principle -- that of having more than one display or display scale. The last depends on the discrimination requirements being much different in one part of a display scale than another, so that the same display presents finely-scaled and coarsely-scaled information in different parts of its range. Figure VII-2 illustrates the five techniques, which are described below.

a. Vernier Displays. Verniers are fine-scale displays used in conjunction with one or more coarser-scaled indicators. When the requirement for fine information is limited to a small portion of the total range of a display, the vernier may refer only to that portion of the range. When the requirement for vernier information extends through much or all of the display range, however, the reference of the vernier must be varied. This can be done by having the vernier represent only the most significant figures of a display, like a micrometer dial, with the vernier reading in, e.g., hundredths and thousandths of an inch, the coarse scale reading in inches and tenths. (Most gas and electric meters have a series of pointer indicators representing successively, e.g., 1000's, 100's, 10's, and units of kilowatt hours, or cubic feet of gas.) An alternative to the presentation of vernier information as the final one or two decimal places of a coarse indicator scale is to reference the vernier to some nominal value that is near the indicated value. A vernier compass for steering a ship, for example, might show on an expanded scale a range of five degrees either side of a nominal value representing the desired course. A vernier display of spacecraft attitude during boost might employ the same technique.

When a vernier display functions throughout the range of a coarse display, the total range number of the two displays is the same as that of a single display with the threshold discrimination of the vernier and the range of the coarse display. This is at most the product of the range numbers of the two displays computed separately.

b. Multiple Element Displays. The multiple element display combines the vernier and coarse scales onto a single instrument, as on a clock or aircraft altimeter. The moving elements are usually pointers, with different pointer shapes defining the appropriate scale divisions. Display range number can be computed much as with the vernier display. There is a danger in such displays of the errors in reading the coarse scale, as when an error of a thousand feet is made in an aircraft altimeter reading -- a sometimes fatal mishap.

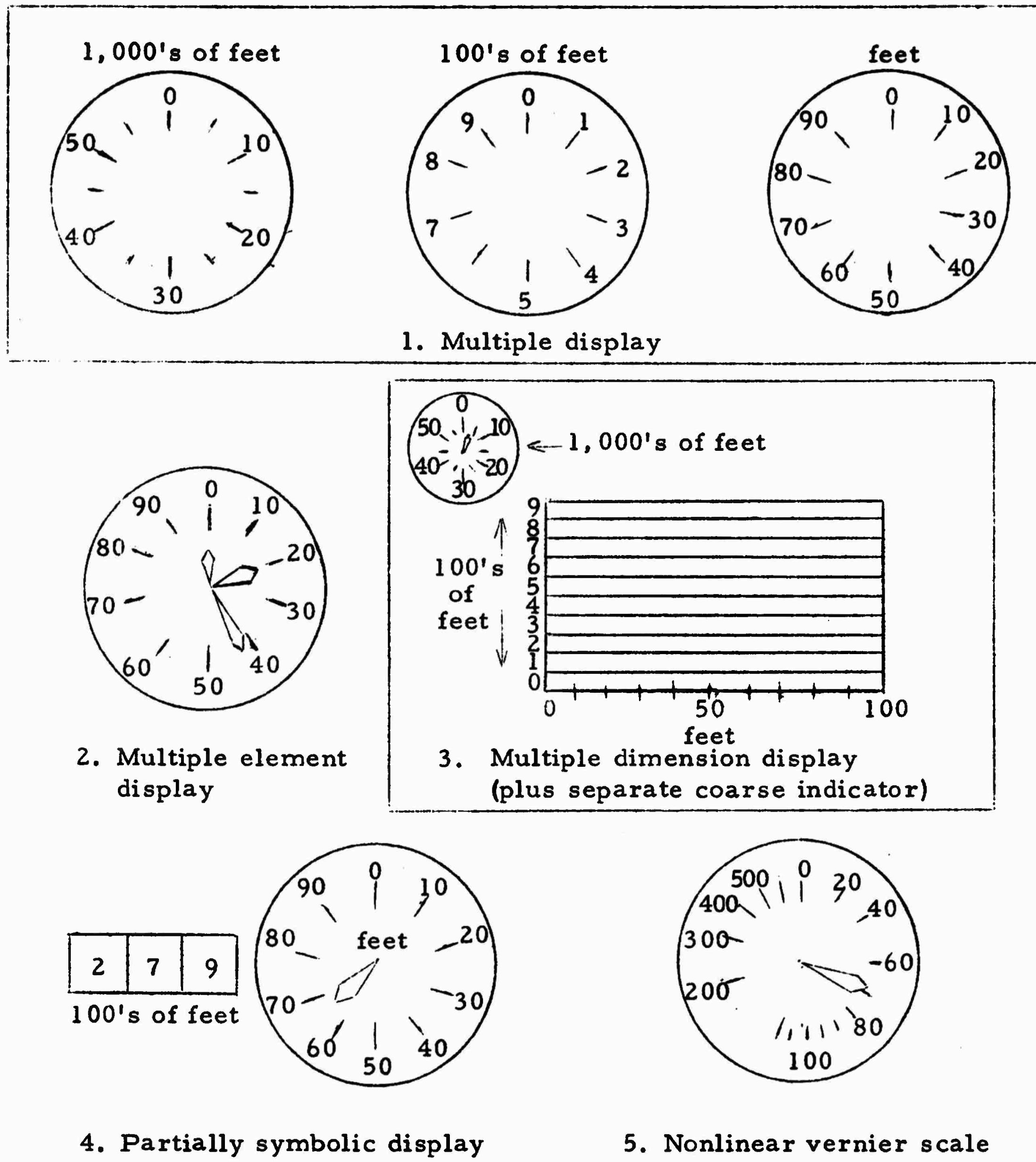


Figure VII-2. Five techniques for expanding the scale of an analog display, illustrated by the aircraft altimeter. (The "nonlinear altimeter" is not considered a practical instrument.)

c. Multiple Dimension Display. Instead of separate display elements, a display may employ separate dimensions to expand its usable scale. A two dimensional display could simply show coarse information on the ordinate and vernier information on the abscissa, for example. The danger in such a presentation is like that of the multiple element display; a large error is possible through misreading the coarse scale.

Dimensions other than spatial dimensions are sometimes employed to expand the scale of displays, as when color (or some other code) is employed to indicate which part of a range is being indicated. The range number is in this case the product of the range number of the display and the number of colors or other coded alternatives.

d. Partially Symbolic Display. The display of coarse information via a counter or other alphanumeric indicator plus the use of a pointer or other analog display of vernier information has much to recommend it. Since the counter is here limited to coarse information, change is slow, while the analog part of the display is not confused with extra pointers; yet the display has the virtues of linear scale and linear motion characteristics throughout the display range. This is often the best solution to the display range problem, and should be adopted more widely than it is. One caution: when the digital counter appears on a pointer dial, as is sometimes necessary, care should be taken not to permit a design in which the pointer can sometimes obscure one digit of the counter.

e. Nonlinear Scales. Often only one part of a display scale must be read with vernier accuracy, and by compressing the scale through nonlinearization, the necessity for a second display or display element can be avoided. Nonlinear scales should be employed with extreme caution, however. Their most serious fault is often not the scale nonlinearity *per se*, but the nonlinearity imparted to the rate of motion of the display element. The operator does not easily adapt to the fact that a given rate of motion of a pointer at one part of a display scale may indicate, e.g., ten times the rate of change of the variable displayed as the same rate of motion of the same pointer when it is at another part of the scale.

C. Display Coordination and Integration

There are at least two sources of information to the operator of a manual control system, e.g., planning information relating to possible or desired outputs, and feedback information relating to the actual output. Frequently there are many sources of information, and more often than not many displays. When there are two or more displays, there is a

problem of display arrangement and coordination, and when a single display combines two or more items of information, there is a problem of display integration.

How displays are coordinated or integrated depends on the relation between or among information signals and the way the information is used by the human operator. Even in simple one dimensional tracking systems, where the operator's task is to match an output to an input signal, there are options as to how the information is presented. These options become more complex when the task is less one of tracking and more one of manual control. This problem is discussed under four headings:

1. Pursuit vs. Compensatory Displays
2. Hierarchical Level and Control Order in Displays
3. Display Arrangement and Coordination
4. Display Integration

1. <u>Pursuit vs. Compensatory Displays</u>

In a tracking task in which the operator matches a continuously variable input signal, the operator will usually require information indicating the input to be matched and the output with which it is matched. In a simple case, the input information may be labelled X_d, the output X. If X_d and X are given to the operator on separate displays or separate elements of one display, the presentation is called "pursuit". The operator employs the control system to pursue or follow X_d with X. If X and X_d are combined and the operator receives only a display of the difference between them (the system error, $X_e = X_d - X$), his display is called compensatory. The operator in this case endeavors to keep X_e small.

The principal advantage of the pursuit display is that by keeping input and system output unmixed, the operator is better able to follow and to anticipate the input. Other things being equal, this permits more accurate tracking except when the input frequency is very low. When the dynamic characteristics of the system are complex, it is especially easy for the operator to confuse the input with results of his own control action unless these are displayed separately.

The statement above, "other things being equal ...." includes, in most comparisons, equal scale. This is unfair, in that a principal advantage of the compensatory display over the pursuit is one of scale.

The compensatory display has a much smaller range number than does the pursuit, and hence can be displayed with equal effectiveness on a smaller display. Pursuit displays show the entire range of the output variable. When the range is large, the magnification or display gain must be small. The compensatory display, on the other hand, need only show the range of the expected error, which may easily be only 10%, 1%, 0.1% or less of the total range of the output. This small range, if expanded to fill the same display, thus results in a display magnification that may be 10, 100, or 1,000 times that of the pursuit display. Such a difference can easlly result in better performance with a compensatory rather than a pursuit display.

Most theoretical research on human tracking performance has utilized compensatory display. The principal reason may be that it is easier to handle theoretically. If the operator receives input signals through two channels and handles them somewhat differently, as he does in the case of pursuit tracking, it is more difficult to analyze what is going on than if there is only one input and one output signal to deal with.

The command display principle can be applied via a compensatory or a pursuit indicator. The pursuit command indicator has two pointers or other display elements, which in a typical application show the ordered as opposed to the actual output of the operator's control. The compensatory indicator would show instead only a single display element, representing the deviation of the operator's control from what is ordered.

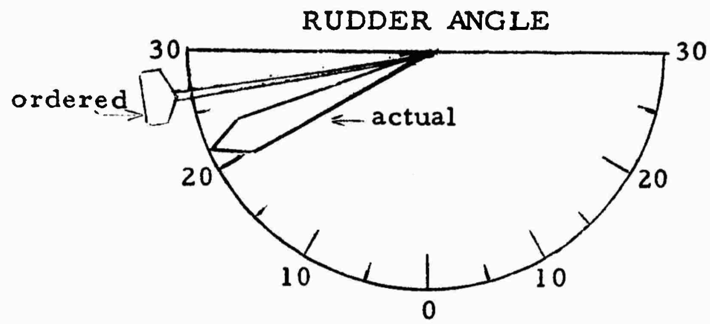
Often it is convenient to display a command signal in terms other than of desired and actual output of a control. Consider the display of command information for steering a large surface ship or submarine. In a manual mode, the operator might turn a helm-wheel to position the ship's rudder, employing a gyrocompass display, a rudder angle indicator, a pit log (speed indicator), and verbal commands ordering the desired heading. In an automatic mode, the automatic steering control system might receive signals representing desired and actual heading, heading rate, and speed, from which an ordered rudder angle signal would be developed. This ordered rudder angle signal would be compared with the actual rudder angle, and the difference employed to open an hydraulic valve moving the rudder to the ordered position.

In a command instrument mode, the helmsman may again order the position of the rudder through the helm wheel.[1] The "automatic" steering

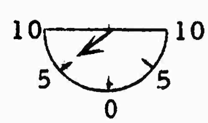
---

[1] On some ships the helmsman orders rudder angle rate rather than rudder angle itself.

control system computes the desired rudder angle from the same information as used in the automatic mode, but this signal is displayed to the operator on a command display. The helmsman now operates the helm to move the rudder to the desired position. A natural display would be of the pursuit type, formed by adding an ordered rudder angle pointer to the rudder angle indicator. Figure VII-3 a. illustrates this command pursuit display. If instead a compensatory display were used, a much smaller display, such as that of Figure VII-3 b., is all that is necessary, as only the difference between desired and actual rudder angles (the two pointers of Figure VII-3 a.) need be shown.



a. Pursuit Indicator (ordered vs. actual rudder angle



b. Compensatory Indicator (rudder angle error)

Figure VII-3. Pursuit and Compensatory command instruments for ship steering.

When the operator has even a little freedom to plan his response, it is usually important to give him a separate display of input or planning information. Consider the rudder angle display of Figure VII-3. Suppose that the operator is allowed some freedom in slowing down or speeding up a correction in heading that is computed for him by his automatic steering system. -- In such a case, it is much easier for him to vary sensibly

from the computed response, i.e., to plan, when he knows both the ordered and the actual rudder angle, and not merely the difference between them.

2. Hierarchical Level and Control Order in Displays

The question of hierarchical level is related so intimately to the nature of displays for manual control that one major aspect of the problem, that of control vs. display augmentation, was presented in Chapter III as part of the treatment of the control system hierarchy, and will not be repeated here (see p. 30). What information is appropriate to an operator in a manual control system depends on the operator's role or hierarchical level. And with the addition of hierarchical levels, the display possibilities multiply.

Consider again the general hierarchical manual control system in which variable X is controlled through changes in Y, which in turn is controlled through changes in Z, with X, Y, and Z all being continuous variables. Z forms the output of the inner loop, and Z_d its input, while Y and X form successively more outward loops. This example might be applied to a variety of control systems, i.e.:

	X	Y	Z	Z_d
1.	Ship heading	rate of change of heading	rudder angle position	helmwheel position (= desired rudder angle)
2.	Wheeled vehicle distance (laterally) off desired path	angle of vehicle to path	front wheel angle	steering wheel angle
3.	X	$\dot{X}$ $(= \frac{dX}{dt})$	$\ddot{X}$ $(= \frac{d\dot{X}}{dt})$	$\ddot{X}_d$

In any case of the three-level hierarchical system of this sort, there are many different fundamental display possibilities. Some of these are described in Table 2. A primary effect of hierarchy in manual control system design is the number and variety of display-control possibilities that multiple hierarchical loops make available to the designer. The choice among such possibilities is often difficult, and the coordination and integration of the displays a problem. There is little systematic research

Table 2

Some Possible Display-Control Variables in Manual Control with a Three-Level Hierarchical System in which Z Causes a Change in Y which Causes a Change in X

Display-Control System Type	Display Computer Inputs	Displays	Human Output	Additional Controller Inputs
Status (Pursuit)	none	$X_d, X, Y,^* Z^*$	Z_d	none
Status (Compensatory)	X_d, X	$X_e, Y,^* Z^*$	Z_d	none
Display Augmented Pursuit	X_d, X	Y_d, Y, Z^*	Z_d	none
Display Augmented Compensatory	X_d, X, Y	Y_e, Z^*	Z_d	none
Command	X_d, X, Y, Z	Z_d	Z_d	none
Control Augmented Pursuit (1)	none	X_d, X, Y	Y_e	Z
Control Augmented Compensatory (1)	X_d, X	X_e, Y	Y_e	Z
Control Augmented Pursuit (2)	none	X_d, X	Y_d	Y, Z
Control Augmented Compensatory (2)	X_d, X	X_e	Y_d	Y, Z
Automatic	none	none required	X_d	X, Y, Z

---

*Display may not be required

to guide the designer. In a paper in 1958 I made the following suggestions for displaying hierarchical information in high-order control systems:[1]

a. There should be an indicator for each integral step between the derivative manipulated by the operator and the system output. Brief exponential lags need not be considered an integral step, even though they increase the order of control.

b. When an integral step in the sequence does not appear on an indicator, the next highest (integral) indicator should be easily differentiable by the operator.

c. Pointer type instruments should be used having linear scales.

d. Indicators should be arranged in derivative sequence with their zero points aligned.

e. Indicator scales should reflect the weights of the corresponding terms in a linear automatic control equation for the system.

f. In vehicular control, output derivatives in space are usually preferable as indicator signals to derivatives in time.

These suggestions still appear sound to me, though explanation is required for some of them. Figure VII-4 illustrates some of these points in a p-nel sketch illustrating submarine depth control. Since the angle of the horizontal rudders (planes) brings about a pitch acceleration in the moving submarine, and pitch results in a change in depth of the moving vehicle, there are four hierarchical levels represented on the panel. Assuming that the displacement of the operator's control corresponds to the position of the planes, suggestions a, c, and d above are followed in the sketch. Since no derivative steps are omitted, suggestion b is inapplicable, although these indicators are easily differentiable by an operator.

---

[1] Kelley, C. R. Instrumentation for Continuous Control. Paper read before the Society of Engineering Psychologists. Washington, D. C.: American Psychological Association, September 1958.

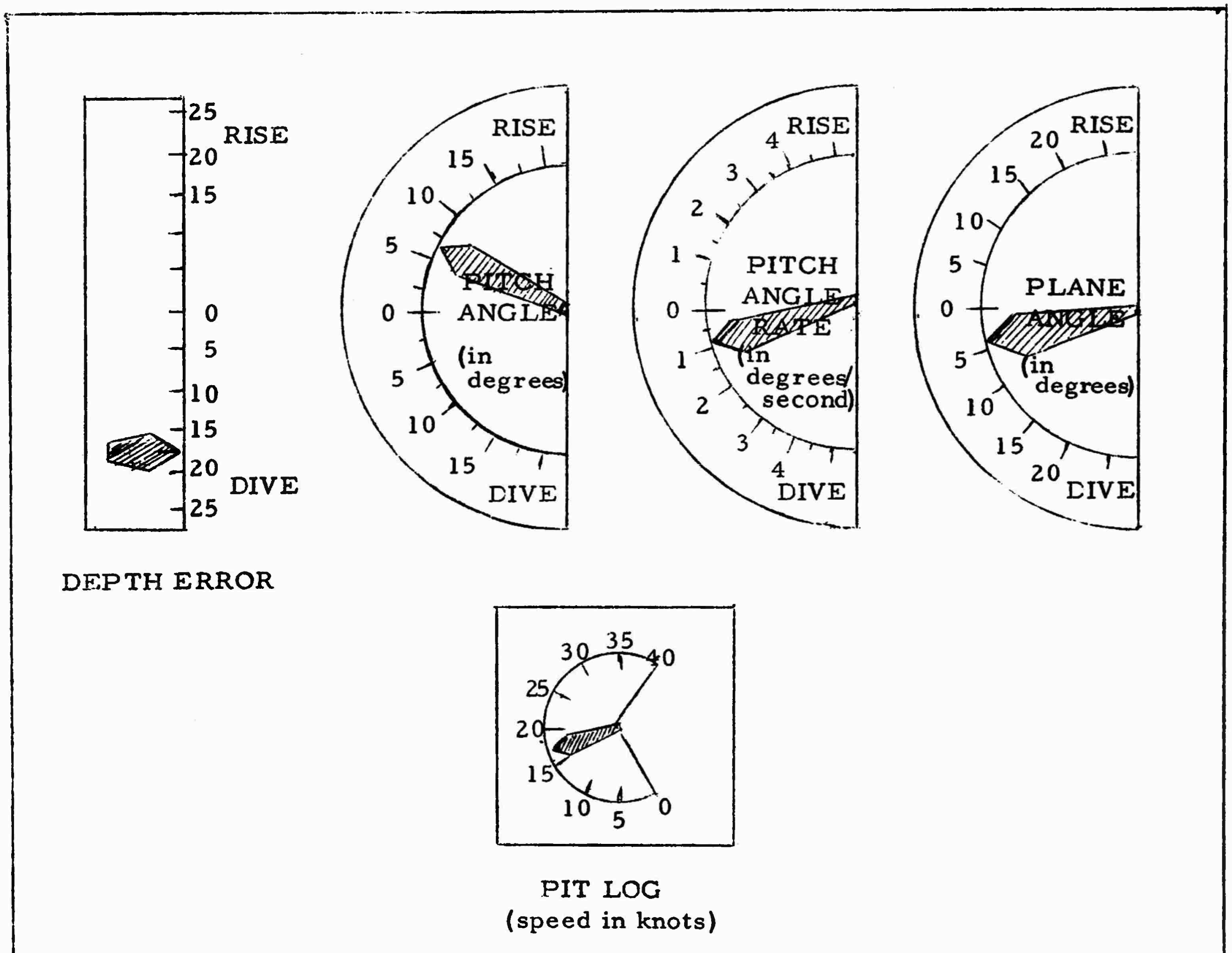


Figure VII-4. Display panel for submarine depth control using existing types of indicators.

Suggestion e concerning scaling on the basis of automatic control equations means simply that it would be desirable if the deviation of any one of the three left hand indicators from zero by a given amount called for an equal corrective deviation in the opposite direction on the right hand indicator, and good control thus be obtained if the operator maintained the average of all deviations at zero. This principle, (which I called "interlinearity") is difficult to achieve in practice, but is nonetheless a useful guide to design. The panel shown does not achieve it for two reasons. First the pitch angle scale is too small in relation to the depth error scale, and for the panel as shown would cause an under-response to pitch angle deviations. Second, the panel includes a time derivative, pitch angle rate, and, to conform to the principle, this would have to change its scale as a function of speed. This brings up suggestion g.

Suggestion g is that space derivatives are usually preferable to time derivatives in vehicle applications. The reason is that time derivatives are affected radically by speed changes, an effect which multiplies with increases in control order. In Figure VII-4, pitch angle is a space derivative function of depth error, and plane angle a space derivative function of pitch angle. To follow a given diving trajectory through the ocean, these indicators will trace approximately the same pattern, irrespective of speed.[1] Pitch angle rate, however, is a time derivative (assumed to be a rate gyro signal) and will vary in proportion to speed, so that the correction called for on the planes for, e. g., a one degree per second error in pitch angle rate, is twice as great when the submarine is travelling at 10 knots as at 20. The panel would be much improved if the pitch rate gyro signal were divided by speed, so that the signal displayed represented rate of change of pitch angle with distance.

3. Display Arrangement and Coordination

The preceding discussion indicates some principles of arranging multiple indicators that are hierarchically related, a subject neglected in the literature. There has been substantial research on principles of arrangement and their application, however. McCormick discusses

---

[1] There is a difference in trajectory due to speed-independent factors affecting control such as pendulosity and fixed plane rates. These differences are slight compared with the changes in time derivative patterns with speed, however.

principles of arrangement by function, by importance, by sequence of use, by frequency of use, and other criteria.[1] He goes on to discuss the application of these principles in particular cases, and the resolution of conflicts between principles. Fogel has an interesting discussion of the arrangement and integration of cockpit instruments.[2] It has been the experience of this author that good arrangements of displays arise from (a) a thorough understanding of the information requirements of the operator; (b) the sensible application of principles and ideas such as those reviewed by McCormick or Fogel or presented by myself in the section above; plus (c) the empirical testing of alternatives via simulation. The application of "principles" without thorough knowledge of the function played by each indicator in the control process can result in very bad arrangements of displays, and almost every novel and difficult display arrangement problem can benefit greatly from simulation.

4. Display Integration

As control systems grow more complex the number of displays grows. The field of human engineering started largely as a result of problems created by the ever-increasing numbers of displays in military aircraft. It has become apparent that for many systems, especially complex vehicles, it is not enough to arrange and coordinate separate displays; something more is required. This "something more" may be supplied by display integration.

The operator of a vehicle builds up an internal model of his vehicle in relation to its environment using both the information on his displays and, when it is available, direct observation. Items of displayed information are sensed by various instruments and displayed. Traditionally the signal from each instrument feeds a separate display; thus the proliferation of instrument dials.

Vehicles like a ship, constrained to move along one body axis on a surface, have few enough degrees of freedom that display proliferation is not a manual control problem. At the opposite extreme is the spacecraft, which may move independently along three position coordinates and around three body axes. Separate displays of position and rate of motion

---

[1] McCormick, E. J. Design and arrangement of controls and displays. Human Engineering, Chapter 14, New York: McGraw-Hill, 1957.

[2] Fogel, L. J. Biotechnology: Concepts and Applications, Chapters 15 and 16. Englewood Cliffs: Prentice-Hall, 1963.

in each degree of freedom thus requires 12 separate indicators. Submarines, fixed wing aircraft, and helicopters fall between these extremes, but all have manual control problems relating to display proliferation. Since the operator of modern complex vehicles invariably has responsibilitie other than vehicle control, the time required to gather and integrate information from many different indicators, as well as the degradation in performance that alternating among indicators causes, are matters of concern.

The operator's internal model of a vehicle under his control is unitary; it does not consist of twelve separate items, but of a single body moving with respect to some environmental reference framework. When display information describing the state of this simple and unitary internal concept must be pieced together from multiple sources, it is obviously more difficult and subject to error than if the information is presented in unitary form. The integration of display information, however, is a serious problem for the engineer. It can easily result in expensive, complex, unwieldy and unreliable display equipment in these early years of the development of display integration techniques. The movement toward integrated displays is gathering force, however, and the engineering problems associated with such displays are being solved. Unfortunately, the same cannot be said of the perceptual and psychological problems involved in the design of such instruments. By all odds the primary problem in the latter category is of display reference. Integrated displays too often attempt to show stationary reference elements as the (apparently) moving elements of a display, causing the problems of motion incompatibility discussed earlier.

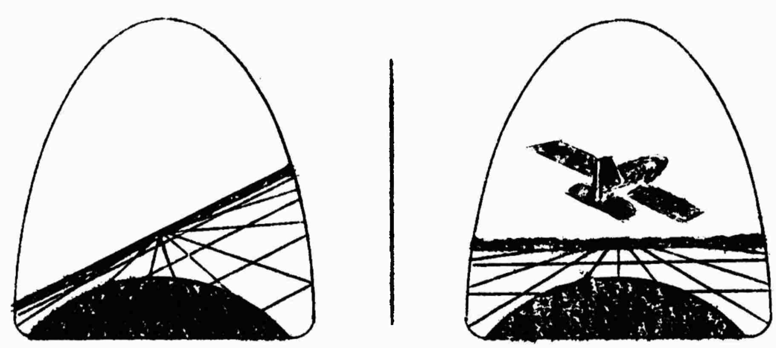
a. Integrated Attitude Displays. An integrated attitude display is one which shows orientation in two or three axes on a single display. An artificial horizon display is integrated, in that it shows both pitch and roll information. Spacecraft since Mercury are turning to a display consisting of a three-axis sphere much like a small globe of the earth which maintains a fixed orientation in inertial space. This display moves in the "wrong way" to the astronaut. Let it be said in defense of the three-axis sphere that despite the motion incompatibility problem, it is an enormous improvement over three separate pointers. Display integration can be a great help, even when incorrectly applied.

The spherical three-axis attitude display need not be motion incompatible. The Hughes Aircraft Company, for example, has developed a laboratory prototype display which is a transparent sphere containing a spacecraft model that moves in three degrees of freedom. This type of display is immediately understandable, is motion compatible, and is easy for even a novice to use.

Motion compatible integrated three-axis attitude displays can be shown on a flat surface, e.g., a CRT tube, by means of a moving vehicle symbol that tilts to show roll, moves up and down to show pitch, and moves left and right to show yaw. This type of display is less simple to interpret than is the spherical spacecraft model display, and attitude rates become distorted on it in a fashion analogous to the problems of the Mercator map projection. Thus at high pitch angles (high "latitudes"), yaw angle rates ("longitudinal" motions) appear greatly magnified.

b. Integrated Attitude Plus Position Displays. The vehicle operator is ordinarily interested, not only in attitude, but also in where he is going, i.e., translational motion. To integrate attitude and translation information becomes a formidable undertaking, especially when the extra degrees of freedom of the helicopter of spacecraft (which need not point in their direction of travel) are dealt with. Whatever display concept is employed, its implementation will not be easy.

One integrated attitude plus translation display of interest has been called the "contact analog".[1] The display concept is of the creation of an analog geometric environment that moves and turns as might a real environment viewed out of the vehicle windscreen (see Figure VII-5a).



a. b.

Figure VII-5. a. Contact analog display, and
b. Tilting airplane display to remedy motion incompatibility problem of contact analog.

---

[1] For examples of description and applications of the "contact analog", see: Integrated Instrumentation for Aircraft (Human Engineering Phase), Report to Douglas Aircraft Company, Calif., by Dunlap and Associates, Inc., December 1954; and Ship Contr 1 VI: Steering and Diving a Submarine with a Contact Analog Display. Groton, Conn.: General Dynamics Corporation, Electric Boat Division Technical Report 411 HF-20, December 1958; and Contact Analog Display. Norden, Division of United Aircraft Corporation, Technical Report TR 0008, January 1962.

This provides an effective technique of integrating the various information items the operator needs. The information is in a form the operator is familiar with, in that it resembles the natural environment in essential features, and so is easy to understand. It does not require special artificial coding of the many items of information that are incorporated into the display, e.g., height, rate of descent or climb, velocity, yaw, pitch, roll, and their changes. However, the contact analog is an "inside out" vehicle coordinating presentation, and research indicates that it does involve motion incompatibility problems and, in consequence, leads to control reversals.[1] By adding a tilting aircraft symbol to show attitude information, i.e., by switching to "outside in" external coordinates for at least roll,[2] the excellent integrative features of the contact analog can be retained, and the motion incompatibility eliminated. The display is, of course, no longer an analogy to contact flight. (See also Figure VIII-3, the integrated predictor display for spacecraft.)

Since the human operator exercises control by reference to an internal model which is not usually one of a fixed vehicle in a moving and twisting environment but of a moving vehicle in a fixed environment, there is good reason to employ displays which show a moving vehicle and fixed environmental reference. As was stated in a previous study by this author:

> There is no reason to believe that the best way to present information about the position and movement of a vehicle to its operator is to show him something in a form akin to what he sees looking out of his vehicle. It would be unfortunate if the potential value of the integrated synthetic display were hurt by unnecessarily clinging to any such limited concept.[1]

Integrated displays are the vehicle displays of the future. It is important at this early stage in their history to analyze and study experimentally all of the various concepts of display integration that exist. We do not know enough to say in advance which are best, for we have only scratched the surface of the possibilities that exist.

---

[1] Kelley, C. R., De Groot, S., & Bowen, H. M. Relative Motion: III. Some Relative Motion Problems in Aviation. Port Washington, N. Y.: U.S. Naval Training Device Center Technical Report 316-2, January 1961.

[2] For a large contact analog display, the "inside-out" vehicle coordinate reference for pitch and yaw is satisfactory, the motion confusion being limited to roll. Ibid.

VIII. SPECIAL DISPLAY TECHNIQUES

There are too many special techniques of display to try to describe even a large minority of them in a one chapter treatment. Only an abbreviated coverage of three topics will be attempted here. The first two of these, historical and predictor displays, consider the effects of freeing the display from the restraint of real time. In the third and final section, the command display technique is more fully described.

A. Historical (Recording) Displays

A primary feature of the internal model of the operator of a control system is that it is free of present time. The model is a result of past experience, and so is an accumulated bit of historical information in itself. Through the model present conditions can be projected into the future. The type of historical display of interest here is recorded status information, the main value of which is to aid prediction.

The perception of motion is an activity that takes place over time, so that the perception of moving displays always involves a short history. At an instant of time, all displays are stationary. The perception of an easily differentiable display signal like a moving pointer, however, is inaccurate compared with the observation of the slope of a line, and while we can hardly perceive time acceleration,[1] we can judge the curvature of a line and even its change in curvature, i.e., its third derivative, quite well. Thus in a high order system, a recording indicator might take the place of four indicators representing position and its first three derivatives and at the same time present the information in a more useful integrated form.[2]

---

[1] Gottsdanker, R. M. The accuracy of prediction motion. Journal of Experimental Psychology, 1952, 43, pp. 26-36.

[2] In 1951 Mr. Herbert Ziebolz of the Askania Regulator Company proposed to the Navy a submarine depth display consisting of a revolving cylinder that continuously traced a depth error record on the front and erased it on the back. I have yet to see such a type of instrument made available, although it appears to be both feasible and needed in many manual control applications.

A historical record is easy to extrapolate forward in time. The "spatialization" of a time-varying quantity makes it possible to take full advantage of the eye's response to space, and to make a spatial extrapolation, i.e., a prediction, in terms of the projected shape of a record trace rather than the projected motion of the moving element of a display. Figure VIII-1 illustrates the point.

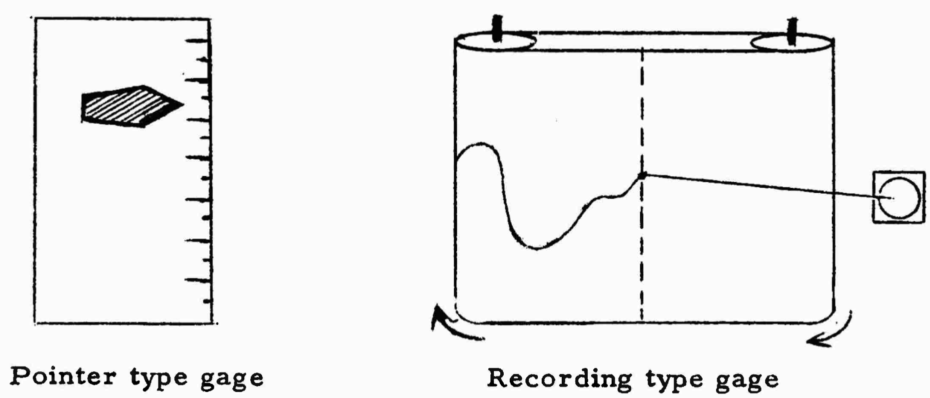


Figure VIII-1. Recording vs. pointer type indicator.

The historical display can be used to better show the relations between or among two or more recorded traces. It converts a relation observable only in terms of separate moving display elements into a spatial relation. This relationship information may be directly useful to the operator in building up his internal model of the system. For this reason, recording displays should have special value in training for manual control, even if they were employed only in training devices.

The operator's internal model operates free of present time, and so may be used to represent past and possible future conditions. The historical display addresses itself to this attribute of the model with an effectiveness exceeded only by the predictor display. It deserves to be utilized more fully in manual control than it has been in the past.

B. Predictor Displays

The historical display allows the operator to better extrapolate past into the future, and so to predict more accurately than does the ordinary

"present time" status display. The predictor display carried this one step further and actually displays one or more predictions for him. Since control is oriented around the future, and the operator employs past and present time information in connection with his internal model to generate predictions, it is no wonder that predictive information, appropriately displayed, can be extraordinarily effective in manual control.

1. The Fast-Time Prediction Technique

This author has been active for many years in the development of predictor displays. He is the inventor of the major technique for generating such displays, the fast-time model method,[1] and has authored or co-authored a series of research studies of the technique.[2] The following description of the fast-time predictor display is drawn verbatim from the paper, "A Predictor Instrument for Manual Control", 1961 revision.

"The predictor instrument provides the operator of a manual control system a display showing information about the predicted future of the variable he is controlling. It does this by means of a special computing device which extrapolates present conditions into the future. Electronic computing means, predominantly of the analog type, have been employed in work to date.

"Ziebolz and Paynter described in 1954 the concept of a two time-scale automatic control circuit in which predictive information about the future of a process controlled was computed.[3] This two time-scale

---

[1] U.S. Patent No. 3,037,201, filed September 2, 1958.

[2] Included among Predictor Instrument Research reports are:
Developing and Testing the Effectiveness of the Predictor Instrument, Office of Naval Research Technical Report 252-60-1. Dunlap and Associates, Inc., March 1960.
A predictor instrument for manual control. In The Predictor Instrument - Final Report and Summary of Project Activities During 1961. Dunlap and Associates, Inc., January 1962.
Predictor instruments look into the future. Control Engineering, 1962, 9(3), pp. 86-90.

[3] Ziebolz, H., & Paynter, F. M. Possibilities of a two time-scale computing system for control and simulation of dynamic systems. Proc. of National Electronic Conference, 1954, 9, pp. 215-223.

concept is of unusual interest. The predictor instrument extends the concept to manual control. The predictive information is generated by an analog of the system to be controlled, operating repetitively on an accelerated time scale. This analog receives signals from sensing instruments responsive to existing conditions in the real system. These signals form the initial conditions with which the analog begins each cycle of accelerated time. The analog system then repetitively computes predictions of the real system's future, which are used to generate one or more displays.

"Figure VIII-2 is a block diagram of the predictor instrument in a manual control loop. The system controlled could be a plane or ship or other vehicle, an elevator, a nuclear reactor -- in fact, any manually controlled dynamic system that responds in a way that can be measured by appropriate instruments, and can be simulated by electronic or other means.

"The heart of the predictor instrument is the fast-time model of the system controlled. This model could be mechanical, electromechanical, or electronic, using either analog or digital methods. We will suppose the fast-time model is a simulation by means of a repetitive electronic analog computer. Sensing instruments in the real system provide signals which are transduced into D. C. voltages and scaled to equal the voltages representing corresponding quantities in the analog model. In this way, the sensing instruments provide initial conditions for the analog system, conditions which begin each cycle of its operation. If the cyclic resetting device resets ["instantly"] 50 times per second and the analog operates on a time scale 500 times that of real time, the analog system will represent the period from present time to 10 seconds (actually 9.98 seconds) into the future. The predictor instrument is completed by using a signal from the output of the analog system to operate an indicator. This indicator presents a signal corresponding to all or part of the prediction period.

"The programmer is a device, usually simple, which represents the assumed control action of the operator during the prediction period. Since the future embraces a range of possible values of the variable controlled, and these are primarily dependent on the control action of the operator, one or a few control actions must be selected and programmed. The most generally useful program has assumed that the operator returns his control to a null position through some appropriate lag. The programmer in such a case may consist of a capacitor discharging through a resistor.

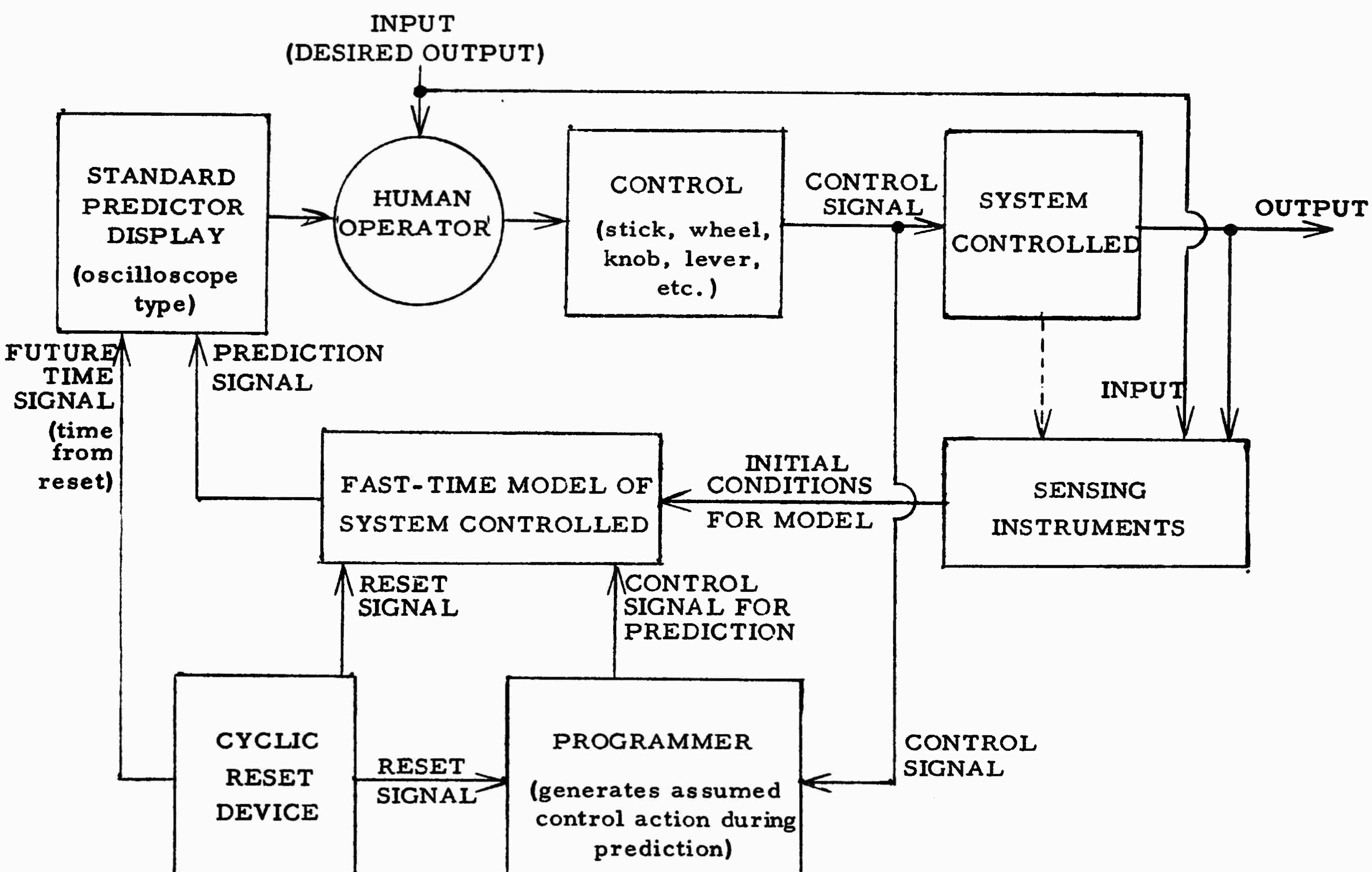


Figure VIII-2. Block diagram of a manual control system utilizing a standard predictor instrument.

"Another very useful program is duplex, assuming that the operator moves his control to either extreme. This second program presents the boundaries of the entire range of possible system performance. In a vehicle, it results in a display of the "maneuvering envelope" within which the vehicle must operate. It is especially useful in enabling an operator to make use of the full maneuvering range available to him. Of course, a realistic maneuvering envelope may be determined by factors other than just the position of the control; e.g., stress on a vehicle or g loadings on an operator. A properly designed predictor instrument could display an envelope described by one or more such factors.

".... Experience indicates that, by using a properly designed predictor instrument, a novice can in 10 minutes or less learn to operate a complex and difficult control system as well as or better than even the most highly skilled operator using standard indicators. The reason for this amazing speed of learning is that the skilled operator with standard instruments must spend weeks or months learning the complex dynamic characteristics of his system so as to be able to know what the system will do under varied conditions. The operator using the predictor instrument knows from the start what his system will do because his instrument tells him; the system's dynamic characteristics are built into the instrument, and in this way are displayed to the operator.

".... The reason the predictor instrument works so well, ... is simply because manual control normally depends on the operator's ability to predict what his system is going to do, and this instrument is the most direct means for giving him the information he requires.

"<u>Predictor vs. Command Instruments</u>.

....Because they are both usually generated by computing circuits and because they both can result in stable, accurate, quickly-learned manual control, there has been a failure to distinguish essential differences between predictor and command instruments. They are entirely different and, in important respects, opposite approaches to manual control instrumentation. When the most desirable system response for all situations can be programmed in advance and when precision and repeatability of response are the principal requirements, the command instrument is a natural choice among manual control instruments. The command instrument is appropriate for the kinds of systems in which an automatic controller is appropriate, but when there are compelling reasons for not going to fully automatic control.

"The predictor instrument ... does not tell the operator what to do, but what he can expect to happen. The operator is in no way "programmed". He may respond quickly or slowly; he may follow a different response trajectory or pattern with time on each recurrence of identical display conditions.

"To illustrate this difference, the operator following a command instrument in diving a submarine will bring the submarine along a predetermined trajectory, which can be followedwith great precision but which cannot be varied without adjusting the instrument in some way; e.g., changing ratios of gains of depth, pitch, and pitch rate signals which make up the command signal. The operator diving a submarine with a predictor instrument is under no constraint as to his trajectory. He may dive slowly, quickly, or at any intermediate rate. The best command instrument will thus result in the most precise and repeatable dive trajectories while the best predictor instrument will result in the most flexible and accurate unprogrammed diving control system."

2. Off-line Control

Many advances on the predictor display technique have been made in recent years. One of the most interesting is the off-line control technique.[1] In off-line predictor systems, the operator tries out control responses on the fast-time model by reference to a predictor display. The response being tried on the model is stored for possible application to the real-time system. When the response is found which brings about the prediction desired, a switch enters it into the real-time system; it is then carried out in real time just as it was tried out on the model.

This technique improves the predictor display's usefulness for planning, because the operator can use it to foresee the results of various possible control alternatives, not just one or a few that are built into the usual predictor display. The off-line technique is undergoing further development at present.

---

[1] Kelley, C. R., Mitchell, M. B., & Strudwick, P. H. Applications of The Predictor Display to Manned Space Flight. National Aeronautics and Space Administration, Office of Manned Space Flight, Technical Report. Santa Monica, Calif.: Dunlap and Associates, April 1964.

3. The Point Prediction Technique

There are other means of generating predictive displays than the fast-time technique, one of the main ones being the "point prediction", which generates a prediction for some fixed period in the future by a real-time computation. To compare the two prediction methods, consider the prediction of the position or angle, Θ, of an undamped body, assuming application of a constant acceleration along a straight line or around an axis. The prediction equation for $\Theta(\tau)$, i.e., Θ as a function of (future) time is:

$$\Theta(\tau) = \Theta + \int_0^{\tau} \dot{\Theta}\, dt + \int_0^{\tau}\int_0^{\tau} \ddot{\Theta}\, dtdt$$

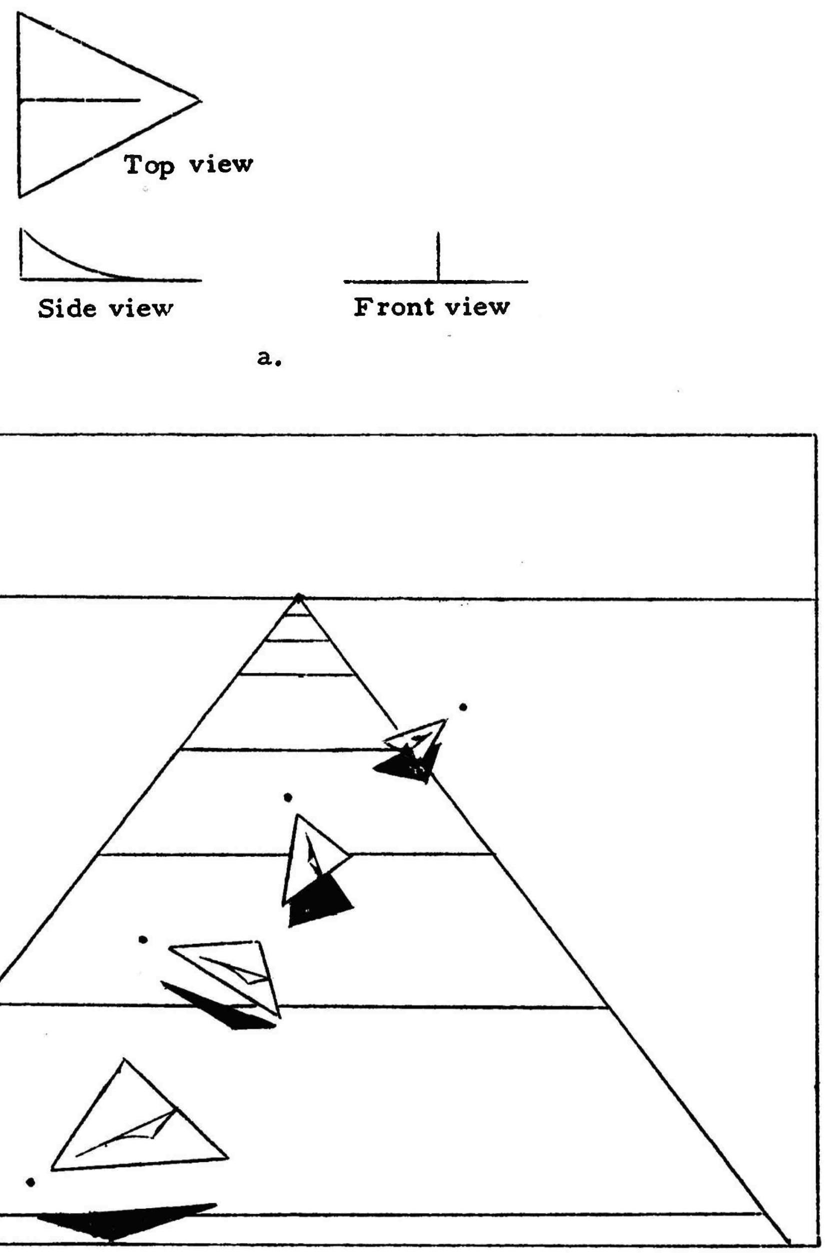
where Θ, $\dot{\Theta}$, and $\ddot{\Theta}$ are the present time (t = o) position, rate, and acceleration terms, respectively. This prediction equation could be mechanized via analog computer. The fast-time model would consist of two integrators cascaded, with $\ddot{\Theta}$ the input, and $\dot{\Theta}$ and Θ, respectively, initial conditions. Repetitive cycling of the integrators would produce repetitive predictions of $\Theta(\tau)$. Each prediction would be continuous from t = o to the end point, a prediction span equal to the cycle length multiplied by the time acceleration ratio.

For a fixed point prediction the equation for Θ is rewritten in the equivalent form,

$$\Theta = \Theta + \tau\dot{\Theta} + \frac{\tau^2}{2}\ddot{\Theta}$$

For any fixed value of τ, i.e., the prediction of Θ at any given future point, T, the equation can be solved by a simple summing amplifier, with the three input signals and coefficients. Several such points could be supplied by several such amplifiers. Note that these points are being predicted continuously in real time, rather than cyclically in fast time. They result in a display of discrete predicted points rather than of a continuous path. Point predictions become more difficult to mechanize for complex equations where there may be, e.g., angles to resolve.

Point predictions can be also obtained from a fast-time predictive model by sampling the model output at the appropriate points, of course, so that a display such as the integrated spacecraft predictor visualized in Figure VIII-3 might be generated by either four point predictors, or one fast-time predictor, the output of which was sampled four times each cycle.



a.

b.

Figure VIII-3. Universal Spacecraft Predictor Display. Depicts predicted position and attitude of a tumbling vehicle with respect to a command path (desired trajectory). Vehicle is at first upside-down.

4. Adaptation via Predictor Display

A principal virtue of the human operator is his ability to adapt to changes in the system under control or in its environment. As was noted in the last chapter, adaptation in manual control involves a change in the operator's internal model. Such changes take place as a result of the operator noting errors in prediction, and adjusting his model accordingly.

Operator adaptation is facilitated by use of a predictor instrument. Errors in prediction are easier to observe when predictions are explicit. One effective form of adaptation is therefore to have manual adjustments of the model which the operator uses to reduce errors in prediction, and so to adapt his display.

This technique can even be used to sense changes in the control system or environment that require remedy. Thus the operator adjusts his model until it is predicting accurately, and the model provides him information that is used to adjust, trim, or realign his real system. An example has been given as to how this technique would be used to detect and correct submarine deviations from neutral buoyancy and trim.[1]

For the finest adjustment in a predictive model, some kind of records of both predicted and actual system behavior is desirable. The combination of historical and predictive display would provide the best form of "adaptive display".

C. Command Displays

The command display does not tell the operator what is happening, but instead tells him what to do. The basic information shown on the command display is not the state of something, but an ordered action. The command instrument says, in effect, "move your control to this position". The operator need not have any idea as to why this action should be taken, but he knows that if he takes it, he will maintain stable control in accord with the precalculated combination of status signals fed into the instrument.

---

[1] Kelley, C. R. A predictor instrument for manual control. In The Predictor Instrument - Final Report and Summary of Project Activities During 1961. Stamford, Conn.: Dunlap and Associates, Inc., January 1962, pp. 15-16.

The operator using a command instrument is in effect given the output of an automatic controller to follow. Design techniques appropriate to automatic control can be applied to develop a command instrument signal. Command instruments are appropriate for systems in which it would be desirable in principle to replace the operator with an automatic system, but where other considerations make it necessary or desirable to keep him in the control loop.

Command instruments bring about extreme accuracy of response in following a precomputed trajectory, and they involve an absolute minimum of learning. When it is desirable to constrain the operator's response to obtain a precise precomputed output, the command instrument is the most effective form of display.

1. History and Theory of the Command Display

The history of the command instrument has not been adequately traced. The author has heard, but has been unable to confirm, that the principle of combining error and error derivative information into a single signal for manual control dates back to fire control systems at or shortly after World War I. The technique was known among control engineers when the writer entered the manual control field in 1951, at which time there were command displays in existence for aircraft flight path indication, ship steering, and submarine diving control. In the latter, the display signal was formed by automatic steering and submarine diving control equipment, and displayed to the operator for his use in an optional "semi-automatic" manual mode of operation.

In 1950, Hick and Bates had diagrammed a display system which summed system error and its first derivative under the title "display aiding". They stated that the aided display control system

> .... consists in adding to the observed misalignment to be minimised a component of rate of change of misalignment. It is indicated in cases in which the control is of the velocity or higher order, because, as will be shown later, there are theoretical and experimental reasons for thinking that anything which reduces the effective order of the control -- i.e., brings it nearer to the positional type -- makes tracking easier ....[1]

---

[1] Hick, W. E., & Bates, J. A. V. The Human Operator of Control Mechanisms, Permanent Records of Research and Development No. 17-204. London, England: Ministry of Supply, May 1950, pp. 10-11.

In 1951 and 1952, this author developed and systematically applied various possible command and partial command display techniques to submarine diving and steering, describing the many different display and control systems possible with different levels of display and control augmentation. This work was released for publication by the Navy in 1953.[1] This may be the first systematic application of the command display technique.

In 1954 the classic report by Birmingham and Taylor of the Engineering Psychology Laboratory of the U. S. Naval Research Laboratory appeared.[2] Birmingham and Taylor christened the technique of the command display "quickening", and since the appearance of this report their Branch of Naval Research Laboratory has been in the forefront of developments in command instrumentation.[3] In their original report, Birmingham and Taylor made a case for having the human operator in a control system do no more than is necessary of those complex operations

---

[1] Kelley, C. R. Submarine Control by a Single Operator. Port Washington, N. Y.: U.S. Navy Special Devices Center Technical Report 954-00-18, October 1953.

[2] Birmingham, H. P., & Taylor, F. V. A design philosophy for man-machine control systems. Proc. of the Institute of Radio Engineers, December 1954, 42(12). Reprinted in Sinaiko, H. W. (Ed.) Selected Papers on Human Factors in the Design and Use of Control Systems. New York: Dover, 1961, p. 75.

[3] See for example:
Birmingham, H. P. et al. A Demonstration of the Effects of Quickening in Multiple-Coordinate Control Tasks. Washington, D. C.: U.S. Naval Research Lab., Report 4380, June 1954.
Taylor, F. V., & Birmingham, H. P. Simplifying the pilot's task through display quickening. Journal of Aviation Medicine, 1956, 27, pp. 27-31.
Sweeney, J. S., et al. Comparative Evaluation of Three Approaches to Helicopter Instrumentation for Hovering Flight. Washington, D. C.: U.S. Naval Research Lab., Report 4954, 1957.
Birmingham, H. P. The optimization of man-machine control systems. IRE Western Electronic Show and Convention Record, Part 4, 1958, pp. 272-276.
Birmingham, H. P., & Taylor, F. V. Why Quickening Works. American Society of Mechanical Engineers Paper 58-AV-9, 1958.
Perry, B. L., & Birmingham, H. P. An analytical method of determining feedback gains for manual control. IRE International Congress on Human Factors in Electronics, May 1962.

that can be done more accurately and quickly by mechanical components.

> .... when a man-machine system must integrate, differentiate, or perform other higher-order computations, these should be supplied by the nonhuman components of the system whenever possible. This is tantamount to saying that the human should be required to do no more than operate as a simple amplifier.[1]

When the criteria for control are strictly those of automatic control, i.e., stability, speed and accuracy of response to a defined input signal spectrum, then the Birmingham and Taylor principle is reasonable. It is not a principle to be applied indiscriminately, however. Why should the human operator necessarily be employed as a "simple amplifier" when he is such a terrible amplifier? What kind of amplifier has a frequency response less than three cycles per second and a transmission delay of more than 200 milliseconds?

The most valuable attribute of the human operator in manual control is his unique ability to plan, to foresee possibilities and choose among them, and not to amplify a display signal. Other displays are far more effective for planning than is the command display. Nonetheless, the command display has proved its value in many situations. It is frequently the cheapest and simplest way to stabilize or add precision to a high-order manual control task. It should be one -- but only one -- of the technques which the designer of the manual control system has available to use when called for.

2. Design of the Command Display

Design of a command display is, in essence, design of the compensation required in an automatic controller for the same system. For sophisticated design, it may require, in addition, compensation for the transmission characteristics of the operator. These characteristics are discussed in some detail in the Appendix. The techniques applicable to automatic controller design can thus be employed in developing the command instrument signal.

---

[1] Birmingham and Taylor, op. cit., 1954.

3. Displaying Command Information

There are many ways in which command information can be displayed. The "straight" command display is usually a rather simple display problem. The requirement is for a display of ordered position of a control of either pursuit or compensatory type. It should, of course, have the same number of degrees of freedom as the control, and with compatible motion relationships, as these have been previously discussed. Unless there are good reasons to do otherwise, the pursuit presentation (separate ordered and actual display elements) will be preferred. If more than one dimension is involved, i.e., if it is a two or three degree of freedom control task, an integrated display is preferable to discrete indicators.

It is actually unusual for an operator to be used solely as a command link. Usually he is at least monitoring system status as well as tracking a command indicator. The combination of command and status information can be a difficult and challenging display design problem. It can be solved in such a way as to give the operator status information and some freedom to plan, so that he has some combination of the virtues of the status instrument for prediction and planning and the command instrument for precomputing a good response trajectory.

4. The Output Plus Command Display

A principal criticism of command displays is that they focus the operator's attention on the innermost loop of the control process, i.e., control position, when he should be thinking in terms of system output and other longer-range outer loop processes. One proposal to remedy this is to add the command signal into the display of output, so that the operator has command information and output information on the same display. This can be done in more than one way, but a two element (pursuit) display consisting of output on one display and a correctly scaled "command plus output" signal on the other is a good form of presentation. The operator can command any output he wishes by tracking the "command plus output" signal to the desired point on the output display and holding it there. The distance between the "command plus output" and actual pointers is then equal to the output error signal, and as long as the actual pointer deviates from the command, corrective action is summed into the "command plus output" display in such a way that in order to keep the display at the desired output, the operator is forced to take the corrective action computed for him. He must put in exactly the control response that will bring the actual output to the command output along the command trajectory.

To illustrate, consider a command display system for control of X, in which desired control position is computed as per this equation:

$$\Delta_d = A(X_d - X) + BY + CZ$$

where Δ_d = the command signal, i.e., the ordered position of the operator's control

X_d and X = desired and actual system output, respectively;

Y and Z are higher order functions of X necessary for stable control, and A, B, and C are fixed coefficients.

A simple command display would display Δ_d and Δ, the desired and actual control position via pursuit command display, or $\Delta_e (= \Delta_d - \Delta)$, a compensatory command display. In either case, the operator would not be required to pay attention to X, his system output. With the "command plus output" display described above, however, the operator would observe a display of X and a display of $(X_d - K\Delta_e)$, i.e., his desired output plus the compensatory command signal, scaled by the coefficient K. To hold this display at X_d, the operator must keep Δ_e at zero. His display appears in terms of the output, however, and while he may not understand the various movements of his control that are necessary to keep the display at X_d, he does think in terms of system output, and can observe how it is carried out.

The scaling requirement for the display is that the output term in the command equation for Δ_e match the desired output term in the equation for the display signal. The command signal is

$$\Delta_e = \Delta_d - \Delta = AX_d - AX + BY + CZ - \Delta$$

D, the signal displayed, however, is

$$D = X_d - K\Delta_e$$

To properly scale Δ_e for this display, it must be true that

$$K = \frac{1}{A}$$

With this the case, the display signal is

$$\begin{aligned} D &= X_d - \frac{1}{A}\Delta_e \\ &= X_d - \frac{1}{A}(AX_d - AX + BY + CZ - \Delta) \\ &= X - \frac{1}{A}(BY + CZ - \Delta) \end{aligned}$$

which is the actual signal employed to generate the display, the X_d terms having cancelled.

5. <u>The Partial Command Display (Display Augmentation)</u>

The technique for developing a partial command (or partially "quickened") display signal is described in Chapter III in the discussion of display vs. control augmentation, and will only be briefly reviewed here. The display is formed by a technique which imposes one or more fixed constraints on how the operator corrects a deviation, but yet does not entirely specify his response. Suppose that an operator of a high order system is required to keep the first derivative of output, $\dot{X}$, proportional to the system error, X_e; this is a typical constraint for partial command systems. The operator could be given a two pointer display of ordered and actual $\dot{X}$, where

$$\dot{X}_d = -AX_e$$

Instead of controlling X, the operator now has the simpler task of controlling $\dot{X}$. The control order of his task is reduced by one. If the further constraint is added that the operator keep the second derivative of output proportion to errors in the first derivative, as defined above, i.e.,

$$\ddot{X}_d = -B\dot{X}_e = -B(\dot{X}_d - \dot{X})$$

$$= -B[-A(X_d - X) - \dot{X}]$$

$$= ABX_d - ABX + B\dot{X}$$

This is then a partial command signal specifying the second derivative of output. The operator's task has been further simplified by a reduction in control order. The display might be a two pointer presentation, the above $\ddot{X}_d$ signal and an actual $\ddot{X}$ signal. Obviously the process could be continued further.

An example should help. Submarine depth, as has been said, is controlled by means of planes which accelerate pitch angle of the moving submarine, which changes depth in consequence of the change in pitch. A partial command display might constrain the ratio of pitch angle to depth error, so that "plus Z" feet of error always constituted an order for "minus Θ" degrees of pitch angle to correct it. The operator then need concern himself only with achieving the desired pitch angle, which would be displayed to him. The process is extended one step by constraining the relation of pitch angle rate to errors in pitch angle. The operator then sees a display of an ordered pitch angle <u>rate</u> signal, formed by combining depth error and pitch angle signals. He matches this order by means of an actual pitch angle rate indicator.[1]

[1] For fuller description, see Kelley, C. R., <u>op. cit.</u>, p. 119.

IX. CONTROLS

After a human operator has performed his internal operations on information received, after he has developed his internal model, perceived the situation, predicted possible outcomes, and planned and programmed his response, the response must be entered into the next stage of his system via a control. Controls are the means by which the operator transmits a signal into a mechanism. We will discuss them under these main headings:

A. The Nature of Human Output

B. The Four Functions of Controls, and

C. Remarks on the Design of Controls.

A. The Nature of Human Output in Manual Control

In control system analysis, it is convenient to specify the output of a link in the system in terms of a variable that can be represented as a single valued function of time. Voltage, temperature, a position or velocity coordinate of a body are such variables. Because the human operator is free of the "present time" constraints of the mechanical system, because he can remember the past and plan for the future, this mode of representation is poorly fitted for describing what man does in a manual control system. The human output is normally a planned action sequence, a series of coordinated movements that are patterned in time and organized about a goal. Even the lower level reflexes are temporal patterns. Any truly adequate means of representing human output must take into account this organization over time. The human operator may respond differently, because of the temporal organization of his response, in situations in which the instantaneous inputs reaching him are the same. It is only in the situation in which man is reduced to a transmission link and does no planning or prediction that his response becomes amenable to analysis by the usual techniques. At this point, and not before, it is reasonable to look for the human output variable that is a single-valued function of time to be measured and analyzed and, if an input can be

similarly described,[1] perhaps entered into an equation for the human operator's transfer function. And at this point it is necessary to inquire whether the human output should be force applied to a control, the motion of the control, or control position.

1. The Human Output Quantity

Accepting the above restrictions, and looking at the operator primarily as a transmission link, it is still difficult to define output of the human, as opposed to that of the control. In operating free-moving knobs and levers, it usually appears to be a position of a control; in handwheel tracking it seems to be a rate, i.e., man keeps the rate of motion of the control approximately equal to the tracking error. In spring-centered controls, force exerted against the control may be a more basic output of the operator than the displacement it brings about. With a "stiff-stick" type of control, force is certainly the basic human output, since there is only a negligible displacement involved.

Some investigators have adopted the point of view that force is the basic human output in manual control, and there is some persuasiveness to the argument. The basic response of our muscles is contraction, and this contraction applies a force tending to bring the two ends of the muscle closer together. This application of force forms the sole basis of human movement, of manipulation and locomotion, and so of control.

The internal process governing manipulation and locomotion, including that involved in even simple tracking type tasks, is not organized around force as such, however, but around the positions and movements of the body and the objects affected by the body. The basic output of a muscle is force, but the basic output of the body in a manual control system is not so clear. For example, the basic output of the postural reflexes would appear to be a position of the body, the basic output of manipulatory control a skilled movement. Surely force must be applied to maintain position in the first case, or to carry out the skilled movement in the second; however,

---

[1] The problem of defining the input mathematically is formidable in manual control save in a few restricted situations. Even when the operator is doing little or no planning, so that he is not originating a response, he is usually reducing information in ways that lend themselves poorly to mathematical analysis. What single-valued function of time can be used as the input to the operator scanning four displays to make a response? What is the input when a man looks at a two-dimensional display, e.g., a time-history or a predicted path, and makes a one-dimensional response? -- And these are simple compared to the case of the operator whose "display" consists of the real world.

the body control process itself at even the lowest level of organization is built around patterns of movement and not contractions of muscles or application of muscular forces *per se*.

As was pointed out in Chapter III, The Control System Hierarchy, what is considered the output of a hierarchical process depends on the level or loop in the hierarchy being discussed; each successively higher loop in the hierarchy has its own output. The level of muscle contraction will have a different output than that of positioning and moving a control. At the latter, higher level the body's output sometimes is a force. Consider applying the brakes in an automobile, where pedal position may vary from one application of the brakes to the other; what is controlled is clearly the force applied. In other tasks, such as adjusting a radio dial or inserting a pin in a socket, position is the basic output, and force may vary from one manipulation to the other. And in the case of handwheel tracking, the rate of motion of crank or handwheel may legitimately be considered the basic output of the operator.

The form of sensory feedback that predominates in a control task seems to determine what appears most "basic" to the operator. Force, velocity, and position feedback cues will usually all be represented in a given control operation, but one may be much stronger. In positioning movements employing visual and tactual position feedback, position is the more "basic" output; in applications of force to a brake pedal, force is the more "basic" output; in spring-centered sticks the case is ambiguous, since position and force cues are both present, but force often predominates over position feedback.

2. Force Operated vs. Position vs. Damped Controls

Whether or not the operator's output is best considered a force, velocity, or position depends, then, on the kind of feedback a control provides. This gives the designer an option, and depending upon the nature and function of the variable to be controlled he will sometimes select one, sometimes another. Except for handwheel or crank tracking, the choice will usually be between "force operated" controls, which might be either spring-restrained or rigid type controls with strain-gage pick-offs, or free-moving "positioning" controls. Some of the important differences between these two types of controls are:[1]

[1] Kelley, C. R. Submarine Control by a Single Operator. Port Washington, N.Y.: U.S. Navy Special Devices Center Technical Report 954-00-18, October 1953.

Force Operated	Position Operated
1. Control forces correspond to forces applied by operator; "natural" control.	1. Control forces do not correspond to forces applied by operator; interpretive step required for control.
2. Control lever is self-centering; forces diminish to zero unless manual force is maintained on the control.	2. Control lever remains at position on which placed; forces remain applied without maintaining manual force on the control.
3. A large range of forces may be accurately controlled in a small range of control lever displacement.	3. To control a large range of forces accurately, a large range of control lever movement is needed.
4. To control a large range of forces accurately, large amounts of manual force are required.	4. A large range of forces can be controlled accurately with very small manual forces.
5. Because of (4), to control a large range of forces accurately, a control must be built and located so the operator may exert large manual forces on it.	5. Because of (4), a large range of forces can be controlled accurately by many types of controls, involving very slight forces, and placed in a large range of locations.

Spring-centered and free positioning controls (which usually maintain a set position by virtue of sliding friction) are not the only choices for controls; wheels and levers for continuous control operation, in particular, may in some cases perform better when viscous damping or inertia are added. Viscously damped controls tend to move at a velocity proportional to the force applied, while high inertia controls move at an acceleration proportional to force applied, and continue moving at a relatively constant rate when force is no longer applied to them. Birmingham and Taylor elaborate on these effects and how they may be applied in system design.[1]

---

[1] Birmingham, H. P., & Taylor, F. V. A design philosophy for man-machine control systems. Proc. of the Institute of Radio Engineers, 1954, 42, pp. 1748-1758.

Howland and Noble[1] tested experimentally the effects of adding inertia, viscous damping, and spring-centering, independently or in combination, in a joystick tracking task. In the task they chose (position tracking of a sinusoidal course) spring-centering improved performance; other possibilities degraded it. It appears from the Birmingham and Taylor work, however, that task conditions could be chosen in which different results obtained. Subsequent work by Chernikoff and Taylor on the interaction between course frequency and control dynamics supports this point of view.[2]

3. Direct vs. Symbolic Output

Just as displays can be classified into symbolic and analog form, human output can be classed into direct and symbolic form. The direct form of output is of primary interest in manual control, wherein an operator pushes a lever, turns a wheel, or operates a pedal that has a direct affect on the course of events.

The symbolic output is a coded message which, like the "message" of the symbolic display, depends for its efficacy on a convention by means of which meanings of an arbitrary sort have been agreed on for the symbols. The symbolic output of an operator may be transmitted to other individuals, and through this communication affect their behavior. This is, of course, a powerful means of control, but is not manual control.

Symbolic controls may also be used to operate equipment, in which the built-in conventional meanings of the symbols are employed by the system designer to facilitate "communication" of man with equipment. The output is in this case a "message" which nobody reads, which is, in the final analysis, only a convenient means of initiating the chain of events required in control. This form of manual control may be simple, as when a dispatcher punches a keyboard number which sends an elevator

---

[1] Howland, D., & Noble, M. E. The effect of physical constants of the control on tracking performance. Journal of Experimental Psychology, 1953, 46, pp. 353-360.

[2] Chernikoff, R., & Taylor, F. V. Effects of course frequency and aided time constant on pursuit and compensatory tracking. Journal of Experimental Psychology, 1957, 53, pp. 285-292.

to the 13th floor. It may also be highly complex, as when the keyboard transmits an instruction to a computer that goes through an elaborate program to derive an output signal which, for example, changes the mixture of chemicals in an automatic manufacturing operation. This latter form of manual control may, in fact, some day be the most important of all. For purposes of this study, however, further discussion will be limited to controls of direct non-symbolic type.

B. The Four Functions of Controls

There are four separate functions played by controls in the manual control process. These have been identified as:[1]

1. Location-identification;
2. Transmission of power;
3. Transmission of information forward from the operator; and
4. Transmission of information back to the operator.

The location and identification of controls involves (primarily) the problem usually discussed under the heading "control coding". The transmission of power concerns the utilization of muscular force in manual control. The third and most important function of a control, however, is the transmission of information, of some form of "message" into a mechanism in the next stage of a man-machine system. The fourth and final function of a control is to feed information back to the operator, i.e., to serve as a form of display.

1. Location and Identification of Controls

The location of a control involves in part the problem of correct identification, as location is one means of coding controls for identification purposes. It involves, in addition, questions as to whether the control can be reached conveniently and operated correctly. The latter problems are in the province of workspace design, and the anthropometric characteristics of the operator population are of central importance. Other characteristics such as the type and range of control motion required, whether or not the operator must exert large forces on the control, the location of associated displays and other controls, and

---

[1] Kelley, C. R. Man and the control process. In Javitz, A. E. (Ed.) Engineering Psychology and Human Factors in Design, Electrotechnology, May 1961.

the sequence of other operations associated with the operation of the given control are all considerations that must be taken into account. Standard texts and handbooks cover these points in some detail.[1]

Control misidentification causes innumerable tragic and unnecessary accidents, and untold inconvenience and annoyance. A classic study of aircraft accidents by Fitts and Jones illustrates the point with respect to cockpit design.[2] We shall never know how many times drivers have turned off their car lights reaching for the heater control or how many housewives have turned on the stove under the wrong pot and left it, with consequences ranging from minor irritation to fatal accidents or fire. We all know from our own experience that the number must be large. It is also largely unnecessary, for the coding of controls has been studied thoroughly, and effective design techniques for minimizing errors in control identification have been developed. For example, studies in shape-coding of knobs by Jenkins[3] and by Hunt[4] are quoted widely, and illustrations derived from these studies appear in most texts and guides treating the design of controls, as do discussions of

---

[1] McCormick, E. J. Human Engineering, Chapters 11-14. New York: McGraw-Hill, 1957.
Ely, J. H., Thompson, R. M., & Orlansky, J. Layout of Workplaces (Chapter V) and Design of Controls (Chapter VI) of the Joint Services Human Engineering Guide to Equipment Design. Ohio: Wright-Patterson Air Force Base, Air Development Center, Technical Reports 56-171 and 172, September and November 1956.

[2] Fitts, P. M., & Jones, R. E. Analysis of Factors Contributing to 460 "Pilot-Error" Experiences in Operating Aircraft Controls. Ohio: Wright-Patterson Air Force Base, Aero Medical Lab. Memorandum Report TSEAA-694-12, 1947. Reprinted in Sinaiko, W. (Ed.) Selected Papers on the Design and Use of Control Systems. New York: Dover, 1961.

[3] Summarized in Jenkins, W. O. Tactual discrimination of shapes for coding aircraft type controls. In Fitts, P. M. (Ed.) Psychological Research on Equipment Design. Washington, D.C.: U.S. Government Printing Office, 1947.

[4] Hunt, D. P. The Coding of Aircraft Controls. Ohio: Wright-Patterson Air Force Base, Air Development Center, Technical Report 53-221, August 1953.

control size, location, color coding, etc.[1]

2. Transmission of Power

Men can produce muscular forces which, for certain systems, are significantly large. These forces are entered into a system by means of a control. Man thus may move the control surfaces on a light aircraft, open a large hydraulic valve to move a ship's rudder, turn the automobile's front wheels in steering and apply the brakes in stopping, each by applying unaided muscular force. As systems get larger and more powerful, the power requirements for such operations tend to increase. The use of means for boosting the power output of man become derirable, if not necessary; witness the use of power brakes and steering on large automobiles. Nonetheless, light vehicles and other systems in which the power output of man can play an important role will stay with us. In addition, many powered controls have manual mechanical back-up systems for reliability and safety. The force with which controls can be operated must therefore remain an important consideration in the design of such systems. Back-up systems usually require more force on a control than is the case with other non-powered controls, and in this respect present a more severe design problem.

Controls requiring substantial muscular force to operate are, like other controls, subject to various accuracy requirements. In other words, the application of force *per se* is only rarely the problem, e.g., with a vehicle emergency brake. More often, the human force output is modulated in time, and frequently with stringent accuracy requirements, as in steering the automobile through traffic or landing the light plane. The control in such cases must allow not only application of substantial muscular force, but the accurate modulation of this force over long periods of time.

---

[1] See in particular:
Gruber, A., *et al*. *Guide to the Coding of Controls*. Technical Guide to the U.S. Army Signal Research and Development Laboratory, Fort Monmouth, New Jersey. Stamford, Conn.: Dunlap and Associates, Inc., January 1962. (An ASTIA Document)
See also:
Ely, J. H., Thompson, R. M., & Orlansky, J. *Design of Controls: Chapter VI of the Joint Services Human Engineering Guide to Equipment Design*. Ohio: Wright-Patterson Air Force Base, Air Development Center, Technical Report 56-172, November 1956.

The locations in which an operator can exert substantial force on a control are few compared with those in which a control requiring little force can be placed. Continuously or frequently operated controls requiring appreciable force in their use require the most preferred positions. Steering wheels, handwheels, joystick, and brake or rudder pedals must occupy the most appropriate positions with respect to the hands and feet of an operator who must exert large forces on them.[1]

The human operator exerting large forces via a control has good information as to what he is doing to the system; feedback information reaches him in an obvious form. This feedback is important in the modulating control movement, and contributes to accuracy and skill. When human operator force requirements are reduced, as they are with powered controls, then the absence of this feedback may become a problem. In this case other techniques of information feedback from the control become appropriate. These are discussed in section 4 below.

3. Transmission of Information via a Control

The primary function of an operator's control in a system employing a source of power outside of man is to transmit information into the system. The operator employs his control to convey a signal, to provide a pattern by means of which the control system will be able to carry out the goal of the control process. The pattern carried by this signal initiates or carries forward the activity of control. It is responsible directly or indirectly for the release of the energies of control in the way required to fulfill the plan of the operator, and that of any higher level of control the operator may be serving.

---

[1] Data on application of force to controls is summarized in:
Ely, Thompson & Orlansky, ibid., and McCormick, E. J., op. cit.
See also:
Wilkie, D. R., Man as a source of mechanical power, Ergonomics, January 1960, 3(1), pp. 1-8.
A classic on its subject is the report by:
Orlansky, J., & Dunlap, J. W. The Human Factor in the Design of Stick and Rudder Controls for Aircraft. Port Washington, N. Y.: U.S. Naval Special Devices Center Report 151-I-8, February 1948.
Human Factors, October 1963, 5(5), is a special issue devoted to biomechanics, with several articles relative to the power output of man.

a. How Information is Carried by a Control. Information is carried as a pattern of energy, a structure in space or in time. This structure can be employed to do something according to a plan, e.g., to open or close a circuit or valve, to modulate a flow of light or electrons or hydraulic fluid. Given the means for control of large amounts of energy by small, the plan may be embodied in the weakest detectable signal. The energy required to code the signal has no relation to the energy the signal can release.

The patterning of control signals in space refers to which of a set of controls is activated and, if there is a range or degree of activation, what the degree, extent or amplitude is. Signals going through a bank of switches or keyboard are coded almost entirely in terms of which controls are activated and when. Continuous controls, on the other hand, require a specification of degree or extent of activation. Signals going through such controls take the form of continuous variables in time.

The patterning of control signals in time is the essence of the human operator's response in manual control. The basic output of the operator of a manual control system was said to be, not a force or rate of motion of a control, but a pattern of movement structured around a plan or goal. Many variations in this pattern are usually possible within the constraints imposed by a given plan. Because the pattern is organized by means of a process that goes beyond the constraints of present time, it must be understood on this basis. The information carried by a control signal is understandable only on the basis of what the operator is striving to bring about.

b. Information Rate. The information rate through a man operated control, particularly through a continuous control, is very small, a few bits per second. The manual control process involves a great reduction of information from sensors to control. Interestingly, the more meaningless the task, the more mechanical the operator performance, the higher his rate of information can be. Typing random combinations of letters or tracking what is to man a high frequency input involves a higher rate of information through controls than does meaningful types of control activity.

There is, of course, a certain minimum rate of information through a control required to, e.g., steer an automobile along a curving road or land an aircraft within satisfactory tolerances. Any task in which the information rate is in excess of man's small information transmission rate must, of course, be mechanized, or designed around the operator in some way. Control augmentation is one important way.

This means automating the inner higher-frequency loop(s) of a control process, but leaving man to perform the longer range, lower-frequency, but lower information rate, outer loop functions. This technique was described in Chapter III, The Control System Hierarchy.

c. Information Significance vs. Information Rate. The amount of information transmitted via a control says nothing, of course, about the importance or usefulness of the information. Information transmitted through a control may be trivial or highly significant or valuable, -- the outgrowth of planning, for example, that could be carried out by no one but a human operator. The control forms the means by which this information is entered into the system. As long as a limiting rate has not been approached, the quantity of information per unit time is a relatively unimportant feature of the control process. The mythical button that would initiate a nuclear attack is, after all, a one-bit control.

d. Scaling and Gain. Controls, like displays, involve problems of scaling, of accuracy or sensitivity vs. range. For accurate control over a variable having a large range, special techniques are also required. A vernier control is one technique whereby, for example, gross positioning or slewing are done by one control, fine adjustments by another. Another technique is to have a second control serve as a gain adjustment or sensitivity switch for the first. A third method is to change to derivative control; when the range of a variable is too great to correspond directly to a control position, the control can instead manipulate that variable's rate of change.

4. Transmission of Information Back to the Operator

Controls not only transmit information forward into the system; they also transmit it back to the operator, closing the inner feedback loops by means of which man regulates and carries out his actions according to plan. The control is also a display, and this feedback aspect of a control is often critical in design. Here the concern is with information transmitted back to an operator through the operation of a control, not with control identification or coding, or the visual features of a control that display its position. With respect to these latter features, the control is, in fact, a display, and the principles of display design can be applied to them. The operational feedback loop from a control, however, presents a form of information to the operator to which vision contributes only secondarily.

The previous discussion of force vs. positioning controls is relevant in this connection. The reason spring-centered controls can bring an improvement in tracking is that they provide better feedback than do free positioning controls. The use of "simulated feel" in aircraft is an extension of the principle involved. "Simulated feel" is a special technique for displaying inner loop feedback information to an operator. In the transition from direct manual operation to powered operation of aircraft control surfaces, it was found that the information feedback from aelerons and elevators, through the control stick to the pilot's hand, was important for fine control of the aircraft. The use of powered control removes this feedback loop, of course, and simulated feel puts back in some of this information artificially.

The principle of simulated feel in aircraft should be employed more in manual control. As control systems get larger and more complex the operator tends to play a role that is increasingly remote from the control process proper, i.e., the application of energy to effect a change in the environment. This lack of immediate contact with the heart of the control process sometimes brings about inaccuracy or unresponsiveness in the system. Adaptations of the "simulated feel" technique can sometimes remedy this indirectness. It can place the operator psychologically in much more direct contact with the environment, and with the means of control. It is one of the major avenues that exists for the improvement of manual control systems.

C. Remarks on the Design of Controls

1. Qualitative Aspects of Control Operations

The controls developed for manual operations are for the most part reflective of the requirements of mechanisms to be controlled, or (too often) of what were the requirements of these mechanisms years or decades ago. There is too often only the most superficial consideration of the characteristics of the human operator. As a result, man is frequently forced into ways of operating that are awkward and limited. His performance reflects, not the manual control capabilities of a man, but the limitations of a machine.

Consider the unimpeded skilled movements of the dancer or acrobat, or the form and quality of human movements where only simple tools or instruments are employed, like the woman knitting or the carpenter using his tools. The qualities of smooth coordination, of rhythm, of movements

that are graceful as well as precise, are characteristic of skilled human performance. In a well designed manual control system, the control movements of an operator exhibit the same qualities.

When a highly practised manual operation is angular or jerky and arhythmic, it is due to the characteristics of the machine. When a vehicle or other controlled body moves unevenly and awkwardly, even when skillfully operated, that, too, is due to the characteristics of the machine. It may be that inherent limitations over which the designer had no control bring about clumsiness of operation. Often it is at least partly due to the fact that the controls, the means by which the operator translates his intentions into actions of the machine, are poorly adapted to such translation.

There is a popular toy on the market that may illustrate the point. It consists of an erasable surface with a writing stylus controlled by two knobs, one of which moves the stylus right and left, the other up and down. By operating the two knobs, one can draw or write with the stylus -- awkwardly, unevenly, imprecisely, slowly. The simplest non-rectangular figure is impossible for the novice, and difficult for the "skilled operator". Would anyone ever use such a clumsy method of manual control over movement in two dimensions in a real control task? The answer is yes, they would, and do. Controls this poorly adapted to human skill are employed in "modern" equipment almost every day. Controls were originally designed around the characteristics of machines of necessity. Now that technology has provided us with techniques for integrating the machine with the man, we too often fail to see the need for -- or even the possibility of -- change. Let me illustrate.

My initial responsibility in the manual control field was to participate as a human factors specialist in the design of the control station for the U.S.S. Albacore, the first large submarine that could be controlled in depth and heading by one man using aircraft-type controls.[1] At this time, diving and steering a large submarine was a seven-man operation,

---

[1] The concept of using aircraft-type controls for submarines was advocated by several individuals in and working with the Navy before I worked on the problem. See:
Trabold, F. W., Tolcott, M. A., & Channel, R. C. Human Factors in the Design of the Submarine Diving Control Station. Office of Naval Research Technical Report SDC 641-1-1. Stamford, Conn.: Dunlap and Associates, Inc., October 1948.

with a man each manipulating the rudder, bow planes and stern planes, and an operator each on the hydraulic, air, and trim manifolds employed to control buoyancy and trim. The seventh man was the diving officer, who coordinated the operations of his six-man team. This arrangement was, as I said at the time, like having seven men drive a car, one steering each front wheel, one each operating clutch, gear shift, accelerator, and brake, and one to "coordinate". -- And all this complexity on the submarine to achieve a result so simple it is done effortlessly and far more skillfully by a fish!

Attitudes toward equipment develop out of necessity, but tend to crystallize and persist to the point when they block improvement. Originally submarines had to be built as they were because of the nature and limitations of the equipment. Once built and operated effectively, their unbelievably awkward manual control arrangement persisted until some pioneer spirits in the Navy took a chance on what seemed to some a radical innovation. In too many manual control systems, the pioneer spirits have not yet been able to make their voices heard.

The above remark about the fish is only partly facetious. The movement of the fish has the quality of integrated skilled motion. The designer of the manual control system should strive to integrate the man into the system in such a way that this same quality is evident.

2. Controls of the Future

Technology is moving ahead rapidly in developing new types of controls, although applications are, in many cases, lagging. The designer has available, in addition to improved handwheels, cranks, pedals, joysticks, wheels, and yokes, many new forms of hand control. There are ball controls, for example. Small balls are used in radar acquisition, while large low-friction balls can be used like a two-dimensional handwheel. There is the stiff-stick control, with strain-gage pick-offs, that may be unequalled for high frequency tracking. Manned spacecraft has brought about the development of many new three axis hand controllers, which should find a variety of other manual control applications.

Many exotic and "futuristic" control concepts are yet within or close to the state of the art. Consider pick-offs of eye position. The eye is biologically specialized for tracking; it should have frequency response characteristics much superior to the hand, for example. The Mackworths

have reported an effective optical pick-off device employing a corneal reflection,[1] while others have employed electrical signals that could be transduced into eye position signals. Lockard and Fozard had previously carried out preliminary control experiments showing that the eye could follow a rapid complex motion in two dimensions with a mean tracking error of less than one degree.[2] Since man-operated tracking systems are going to be with us for many years still, we ought to try to utilize for tracking the best tracking element the body has, and this is the eye. I know of no further work being carried out to this end, however.

New developments in bioelectric transducers have made available a whole family of signals of potential value for control. We could, if we wished, control air temperature from the skin temperature of the operator. We could equally well use electromyographic pick-offs from various muscles to use for such things as vehicle control or tracking. A proper selection of muscles might, for example, provide integrated three-axis control with less cross-coupling of axes than is present in a three-axis hand controller. Such a system would have obvious advantages for an operator inside a heavy pressure suit.

The technique of entering the control response into a computer which processes it, applying logical operations and stored information, and initiates the control action, has been barely mentioned, and cannot be done justice here. That would require a book in itself.

We can just begin to envisage the possibilities opening up in the field of man-operated controls and control devices. The field of manual control began with the first hominoid to use a tool. -- Yet it appears to just be getting under way.

---

[1] Mackworth, J. F., & Mackworth, N. H. Eye fixations recorded on changing visual scenes by the television eye marker. Journal of the Optical Society of America, July 1958, 48(7), pp. 439-445.

[2] Lockard, R. B., & Fozard, J. L. The Eye as a Control Mechanism. China Lake, Calif.: Naval Ordnance Test Station, Technical Report 1546, August 1956.